

Optimalizace

Elektronická skripta předmětu B0B33OPT,
verze ze **18. ledna 2022.**

Tomáš Werner
werner@fel.cvut.cz



Katedra kybernetiky
Fakulta elektrotechnická
České vysoké učení technické

Obsah

1	Značení a základní pojmy	1
1.1	Značení	1
1.1.1	Množiny	1
1.1.2	Zobrazení	2
1.1.3	Funkce a zobrazení více reálných proměnných	3
1.2	Extrémy funkce na množině	4
1.3	Úloha spojité optimalizace	5
1.4	Cvičení	13
I	Použití lineární algebry v optimalizaci	16
2	Maticová algebra	17
2.1	Operace s maticemi	18
2.1.1	Sčítání a násobení skalárem	18
2.1.2	Maticový součin	18
2.1.3	Transpozice	18
2.1.4	Inverze	19
2.1.5	Determinant	19
2.1.6	Stopa	20
2.1.7	Standardní skalární součin a norma	20
2.2	Matice s jedním sloupcem nebo jedním řádkem	21
2.3	Matice sestavené z bloků	22
2.4	Co je soustava lineárních rovnic?	23
2.5	Maticové zločiny	25
2.6	Cvičení	27
3	Linearita	32
3.1	Podprostory	32
3.1.1	Komplementární podprostory a projekce	34
3.2	Lineární zobrazení	35
3.2.1	Prostor obrazů a nulový prostor	37
3.2.2	Hodnota	38
3.2.3	Matice s plnou hodnotou	39
3.2.4	Věta o hodnotě a nulitě	40
3.2.5	Projekce	41
3.2.6	Zrcadlení a reflektory	43

3.3	Afinní podprostor a zobrazení	44
3.4	Cvičení	48
4	Ortogonalita	52
4.1	Délky, úhly, vzdálenosti	52
4.2	Ortogonální a ortonormální vektory	52
4.3	Matice s ortonormálními sloupci	53
4.4	Gram-Schmidtova ortonormalizace	55
4.4.1	QR rozklad	56
4.5	Ortogonální podprostory	57
4.5.1	Vztah k prostoru obrazů a nulovému prostoru	58
4.6	Ortogonální projekce a projektoři	59
4.6.1	Ortogonální zrcadlení a reflektory	60
4.6.2	Vzdálenost bodu od podprostoru	60
4.7	Cvičení	62
5	Nehomogenní lineární soustavy	66
5.1	Přibližné řešení ve smyslu nejmenších čtverců	66
5.1.1	Ortogonální projekce na podprostor daný obecnou bází	68
5.1.2	Řešení pomocí QR rozkladu	68
5.1.3	Lineární regrese	69
5.1.4	Statistické odůvodnění kritéria nejmenších čtverců	70
5.1.5	Vícekritériální nejmenší čtverce, regularizace	71
5.2	Řešení s nejmenší normou	71
5.3	Pseudoinverze obecné matice	73
5.4	Cvičení	74
6	Spektrální rozklad a kvadratické funkce	78
6.1	Vlastní čísla a vektory	78
6.1.1	Spektrální rozklad symetrické matice	81
6.2	Kvadratická forma	83
6.2.1	Definitnost kvadratické formy / její matice	84
6.2.2	Definitnost ze znamének hlavních minorů	85
6.2.3	Diagonalizace kvadratické formy	85
6.2.4	Choleského rozklad	87
6.3	Kvadratická funkce	87
6.3.1	Doplnění na čtverec	88
6.3.2	Kvadrika	89
6.4	Cvičení	90
7	PCA a SVD	95
7.1	Úloha na nejmenší stopu	95
7.2	Proložení bodů podprostorem	97
7.2.1	Jiný pohled: nejbližší matice nižší hodnosti	99
7.2.2	Co když prokládáme afinním podprostorem?	99
7.3	Přeuročené homogenní lineární soustavy	100
7.4	Singulární rozklad (SVD)	102
7.4.1	Nejbližší matice nižší hodnosti z SVD	103

7.4.2	Extrémy lineární funkce isometrie	104
7.5	Cvičení	105

II Nelineární optimalizace 109

8	Derivace	110
8.1	Nelineární zobrazení	110
8.2	Spojitosť	111
8.3	Derivace funkce jedné proměnné	112
8.4	Parciální derivace	113
8.5	Totální derivace	114
8.5.1	Derivace složeného zobrazení	115
8.5.2	Maticový kalkulus	116
8.6	Směrová derivace	119
8.7	Gradient	120
8.8	Parciální derivace druhého řádu	121
8.9	Shrnutí značení derivací	122
8.10	Taylorův polynom	122
8.11	Cvičení	124
9	Extrémy funkce na množině	127
9.1	Vnitřek a hranice množiny	127
9.2	Existence globálních extrémů	128
9.3	Lokální extrémy	130
9.4	Cvičení	131
10	Volné lokální extrémy	134
10.1	Analytické podmínky	134
10.2	Iterační metody na volné lokální extrémy	136
10.2.1	Volba délky kroku	137
10.3	Gradientní metoda	138
10.3.1	Závislost na lineární transformaci souřadnic	138
10.4	Newtonova metoda	139
10.4.1	Použití na soustavy nelineárních rovnic	139
10.4.2	Použití na minimalizaci funkce	140
10.5	Nelineární metoda nejmenších čtverců	141
10.5.1	Gauss-Newtonova metoda	141
10.5.2	Rozdíl oproti Newtonově metodě	143
10.5.3	Levenberg-Marquardtova metoda	144
10.6	Cvičení	144
11	Lokální extrémy vázané rovnostmi	148
11.1	Lineární omezení	149
11.2	Nelineární omezení	151
11.2.1	Tečný prostor	152
11.2.2	Podmínka prvního řádu	154
11.2.3	Podmínky druhého řádu	156

11.3 Cvičení	157
III Lineární programování	161
12 Lineární programování	162
12.1 Transformace úloh LP	163
12.1.1 Po částech afinní funkce	164
12.2 Jednoduché úlohy LP	165
12.3 Typické aplikace LP	167
12.3.1 Optimální výrobní program	167
12.3.2 Směšovací (výživová) úloha	167
12.3.3 Dopravní úloha	168
12.3.4 Distribuční úloha	169
12.4 Přeurčené lineární soustavy	169
12.4.1 Vektorové normy	169
12.4.2 Přibližné řešení lineárních soustav v 1-normě a ∞ -normě	170
12.4.3 Lineární regrese	171
12.5 Celočíselné lineární programování, LP relaxace	172
12.5.1 Nejlepší přiřazení	173
12.5.2 Nejmenší vrcholové pokrytí	174
12.5.3 Největší nezávislá množina	175
12.6 Cvičení	176
13 Konvexní množiny a mnohostěny	181
13.1 Konvexní množiny	181
13.2 Čtyři kombinace, čtyři obaly	182
13.3 Konvexní mnohostěny	183
13.3.1 Extremální body	184
13.3.2 Stěny mnohostěnu	186
13.3.3 Extrémy lineární funkce na mnohostěnu	187
13.4 Cvičení	188
14 Simplexová metoda	192
14.1 Stavební kameny algoritmu	194
14.1.1 Přejít k sousední standardní bázi	194
14.1.2 Kdy je nové bázové řešení přípustné?	195
14.1.3 Co když je celý sloupec nekladný?	195
14.1.4 Ekvivalentní úpravy účelového řádku	196
14.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?	196
14.2 Základní algoritmus	197
14.2.1 Cyklení	199
14.3 Inicializace algoritmu	199
14.3.1 Dvoufázová simplexová metoda	200
14.4 Cvičení	202

15 Dualita v lineárním programování	205
15.1 Konstrukce duální úlohy	205
15.2 Věty o dualitě	206
15.3 Příklady na konstrukci a interpretaci duálních úloh	210
15.3.1 Mechanické modely	213
15.4 Cvičení	215
IV Konvexní optimalizace	218
16 Konvexní funkce	219
16.1 Vztah konvexní funkce a konvexní množiny	221
16.2 Konvexita diferencovatelných funkcí	222
16.3 Operace zachovávající konvexitu funkcí	223
16.3.1 Nezáporná lineární kombinace	224
16.3.2 Skládání funkcí	224
16.3.3 Maximum	225
16.4 Cvičení	226
17 Konvexní optimalizační úlohy	229
17.1 Příklady nekonvexních úloh	229
17.2 Konvexní optimalizační úloha ve standardním tvaru	230
17.3 Ekvivalentní transformace úlohy	231
17.4 Třídy konvexních optimalizačních úloh	232
17.4.1 Lineární programování (LP)	232
17.4.2 Kvadratické programování (QP)	232
17.4.3 Kvadratické programování s kvadratickými omezeními (QCQP)	233
17.4.4 Programování na kuželu druhého řádu (SOCP)	234
17.4.5 Semidefinitní programování (SDP)	234
17.5 Konvexní relaxace nekonvexních úloh	235
17.6 Cvičení	236
18 (*) Lagrangeova dualita	238
18.1 Minimální nerovnost	238
18.2 Lagrangeova duální úloha	239
18.3 Silná dualita	240
18.4 Příklady	241
19 Vícekriteriální optimalizace	243
19.1 Uspořádání na množině	243
19.2 Úlohy vícekriteriální optimalizace	244
Rejstřík	245

Kapitola 1

Značení a základní pojmy

Matematické výrazy na zvláštním řádku (rovnice apod.) jsou ve skriptech číslovány číslem v závorce, takže např. (3.1) je vždy odkaz na rovnici. Na matematické věty se odkazujeme jako Věta 3.1, podobně na tvrzení, důsledky a lemata. Odkazy na kapitoly, sekce a podsekce jsou značeny znakem paragraf, např. §3 je odkaz na kapitolu 3 a §3.1 je odkaz na sekci 1 v kapitole 3. Na řešené příklady se odkazujeme jako Příklad 3.1. Na konci každé kapitoly jsou cvičení, na které se odkazujeme slovem Cvičení 3.1. Když tedy napíšeme třeba ‘podívejte se na Příklad 3.1’, znamená to něco *úplně jiného* než ‘podívejte se na Cvičení 3.1’. Pokud potkáte slovo či sousloví vysázené **tučně**, jde o nově zavedený pojem, který máte pochopit a zapamatovat si. Ve skriptech tedy nepotkáte číslované definice, protože všechny tučně vysázené pojmy jsou definice. Všechny tyto pojmy (a nějaké další) jsou v *rejstříku* (a všechny pojmy z rejstříku musíte znát a chápat ke zkoušce!). Slova vysázená *kurzívou* znamenají buď zdůraznění, nebo poprvé zmíněný důležitý avšak všeobecně známý pojem. Odstavce, věty, důkazy, příklady a cvičení označené hvězdičkou (*) jsou rozšiřující (a často zajímavé) ale nejsou nezbytné ke zkoušce. Je-li za nějakým tvrzením v závorce pokyn *odvod'te!*, *proved'te!*, *ověřte!* či otázka *proč?*, znamená to, že správnost tvrzení není patrná na první pohled, ale snadno ho dokážete, když se nad ním zamyslíte nebo si to napíšete na papír (pokud to nesvedete, něco je špatně a musíte se zastavit a studovat příslušnou pasáž znovu).

1.1 Značení

Zopakujme matematické značení, které se používá v celých skriptech a které by student 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	kartézská mocnina množiny, $A^n = \underbrace{A \times \dots \times A}_{n\text{-krát}}$
\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

Desetinná čísla budeme psát anglicky s desetinnou tečkou, např. 1.23 namísto českého 1,23.

1.1.2 Zobrazení

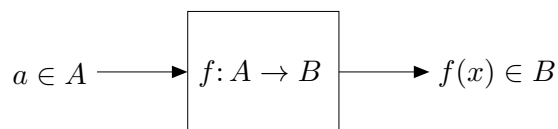
Binární *relace* (angl. *relation*, tedy ‘vztah’) množin A, B je podmnožina jejich kartézského součinu,

$$f \subseteq A \times B. \quad (1.1)$$

Jestliže $(a, b) \in f$ a $(a, b') \in f$ implikuje $b = b'$ (tj. každý prvek množiny A je v relaci s nejvýše jedním prvkem množiny B), pak relaci nazýváme *zobrazení* z množiny A do množiny B a místo (1.1) ho značíme

$$f: A \rightarrow B \quad \text{nebo (méně často)} \quad A \xrightarrow{f} B \quad (1.2)$$

a místo $(a, b) \in f$ píšeme $b = f(a)$. Zobrazení si můžeme představit jako ‘černou skříňku’, která každému prvku $a \in A$ (vzoru) přiřadí právě jeden prvek $b = f(a) \in B$:



Slovo *funkce* (angl. *function*) formálně znamená přesně totéž jako zobrazení (angl. *mapping* or *map*), často se však používá pro zobrazení do číselných množin (tedy $B = \mathbb{R}, \mathbb{Z}, \mathbb{C}$ apod.).

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

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

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

$$f(A') = \{ f(a) \mid a \in A, \varphi(a) \} \quad (1.4)$$

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

Složení dvou zobrazení $f: A \rightarrow B$ a $g: B \rightarrow C$ (přehledněji píšeme $A \xrightarrow{f} B \xrightarrow{g} C$) je zobrazení $g \circ f: A \rightarrow C$ s hodnotami $(g \circ f)(a) = g(f(a))$ pro každé $a \in A$. Skládání je *asociativní*: pro $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$ platí $(h \circ g) \circ f = h \circ (g \circ f)$. Proto můžeme závorky vynechat a psát jen $h \circ g \circ f$. Takto lze definovat složení libovolného počtu zobrazení.

Je-li $A' \subseteq A$, *restrikce* (neboli *zúžení*) zobrazení $f: A \rightarrow B$ na množinu A' je zobrazení $f|_{A'}: A' \rightarrow B$, které má na A' stejné hodnoty jako f , tj. $f|_{A'}(a) = f(a)$ pro každé $a \in A'$.

Zobrazení $f: A \rightarrow B$ se nazývá

- *injektivní* (neboli *prosté*), jestliže každý vzor má jiný obraz, tj. $f(a) = f(a') \Rightarrow a = a'$.
- *surjektivní* (neboli *na*), jestliže každý obraz má aspoň jeden vzor, tj. $f(A) = B$.
- *bijektivní* (neboli *vzájemně jednoznačné*), jestliže je injektivní a surjektivní. K bijektivnímu zobrazení existuje *inverzní zobrazení* $g: B \rightarrow A$ tak, že $f(x) = y \Leftrightarrow x = g(y)$ pro každé $x \in A$ a $y \in B$ (pak píšeme $g = f^{-1}$). Je-li A konečná, bijekce se také nazývá *permutace*.

Transformace je zobrazení nějaké množiny do sebe¹, tedy $f: A \rightarrow A$. Transformace se nazývá

- *idempotentní*, jestliže $f \circ f = f$, tedy pro každé $a \in A$ platí $f(f(a)) = f(a)$.
- *involuce*, jestliže je inverzí sebe sama, tj. pro každé $a \in A$ platí $f(f(a)) = a$.

1.1.3 Funkce a zobrazení více reálných proměnných

Symbol

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-krát}}$$

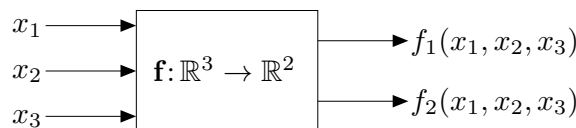
značí kartézský součin množiny \mathbb{R} se sebou n -krát (jak jsme zavedli v §1.1.1. Každý prvek množiny \mathbb{R}^n je tedy uspořádaná n -tice reálných čísel $\mathbf{x} = (x_1, \dots, x_n)$, které budeme říkat (n -rozměrný) **vektor**². Zápis

$$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m \tag{1.5}$$

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 $\mathbf{f} = (f_1, \dots, f_m)$.



¹Ne všichni autoři používají slovo ‘transformace’ v tomto smyslu, je to jen široce používaná konvence spíše než definice. Používá se zejména tehdy, má-li f geometrickou interpretaci.

²I když slovo *vektor* má v lineární algebře obecnější význam jako prvek libovolného (axiomy definovaného) lineárního prostoru, ve skriptech potkáte pouze vektory z \mathbb{R}^n a občas \mathbb{C}^n . Takové vektory se obvykle píší tučně \mathbf{x} spíše než \vec{x} . V pokročilejších textech se vektory značí jednoduše kurzívou x jako skaláry.

Pro $m = 1$ jsou hodnotami zobrazení skaláry a proto budeme v tom případě 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, budeme často pro $m = 1$ mluvit o *funkci* a pro $m > 1$ o *zobrazení*.

Je-li $n \in \{1, 2, 3\}$, často budeme proměnné zobrazení $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ značit x, y, z místo x_1, x_2, x_3 . Např. pro $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ je pohodlnější psát $f(x, y) = x^2 - xy$ místo $f(x_1, x_2) = x_1^2 - x_1x_2$.

1.2 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 . Chceme-li být přesnější, takový prvek x nazýváme *minimalizující argument* nebo *argument minima* (angl. *minimizer*) funkce na množině. Číslo $f(x)$ nazýváme *minimální hodnotou* funkce f na množině X a píšeme

$$f(x) = \min_{x' \in X} f(x'). \quad (1.6)$$

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

$$\operatorname{argmin}_{x \in X} f(x) = \{x \in X \mid f(x) \leq f(x') \forall x' \in X\} \subseteq X. \quad (1.7)$$

Na minimum funkce na množině se dá podívat i poněkud abstraktněji. Necht' Y je nějaká množina reálných čísel, tedy $Y \subseteq \mathbb{R}$. Prvek $y \in Y$ nazveme *nejmenší* (nebo také *minimální*) *prvek* množiny Y , jestliže $y \leq y'$ pro všechna $y' \in Y$. Tento nejmenší prvek značíme $\min Y$. Ne každá množina $Y \subseteq \mathbb{R}$ má nejmenší prvek (např. interval $(0, 1]$ ho nemá). Pokud ho ovšem má, má pouze jeden.

Označme

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

obraz množiny X funkcí f (viz §1.1.2). Pokud množina $f(X)$ má nejmenší prvek, platí tedy

$$\min_{x \in X} f(x) = \min\{f(x) \mid x \in X\} = \min f(X). \quad (1.8)$$

Budeme používat oba zápisy, (1.6) i (1.8).

Funkce nemusí mít na množině minimum, což plyne z toho, že ne každá množina $Y \subseteq \mathbb{R}$ má minimální prvek. V tom případě je množina (1.7) prázdná.

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

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' $(a_1, a_2, \dots, a_5) = (1, 2, 3, 2, 3)$. Pak³ $\max_{i=1}^5 a_i = 3$, $\operatorname{argmax}_{i=1}^5 a_i = \{3, 5\}$. ◆

1.3 Úloha spojitě optimalizace

Matematickou optimalizací rozumíme hledání minima nějaké funkce $f: X' \rightarrow \mathbb{R}$ na nějaké množině $X \subseteq X'$. Tato formulace je velmi obecná, neboť množina X může být zcela libovolná. Obvykle se rozlišují tři široké kategorie úloh:

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

Tento kurs se zabývá *spojitou optimalizací*, ve které $X' = \mathbb{R}^n$ 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

$$g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \quad (1.9a)$$

$$h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l \quad (1.9b)$$

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 jsou obvykle spojitě a často i diferencovatelné a množina X je obvykle nespočetná a souvislá (nebo je aspoň sjednocením máleho počtu souvislých množin). Říkáme také, že minimalizujeme funkci $f(\mathbf{x})$ za podmínek (1.9) příp. (1.11). To lze psát také jako

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmíněk} \quad & g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \\ & h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l \\ & x_1, \dots, x_n \in \mathbb{R} \end{aligned} \quad (1.10)$$

Podmínka $x_1, \dots, x_n \in \mathbb{R}$ není vlastně omezením, často se ale píše, aby bylo jasné, co jsou proměnné úlohy. Toto je tedy *úloha spojitě optimalizace v obecném tvaru*.

Soustavu (1.9) lze napsat kratěji ve vektorovém značení jako

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \quad (1.11a)$$

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}, \quad (1.11b)$$

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. Tedy

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \}. \quad (1.12)$$

³Místo $\max_{i=1}^5 a_i$ se obvykle píše $\max_{i=1, \dots, 5} a_i$. Používáme zde první způsob v analogii se značením $\sum_{i=1}^5 a_i$.

Ve shodě se značením (1.4) se pak úloha (1.10) může zapsat jako

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}. \quad (1.13)$$

V matematické analýze se řešením úlohy (1.10) říká *extrémy funkce f vázané podmínkami (1.9)*. 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, kriteriální) funkce.
- Prvky množiny X se nazývají *přípustná řešení*, což je vlastně protimluv, protože nemusí být řešeními úlohy (1.10). Prvkům množiny $\operatorname{argmin}_{\mathbf{x} \in X} f(\mathbf{x})$ se pak říká *optimální řešení* nebo krátce *optima* úlohy. Číslu $\min_{\mathbf{x} \in X} f(\mathbf{x})$ se říká *optimální hodnota* úlohy.
- Rovnice a nerovnice (1.9) se nazývají *omezující podmínky*, krátce *omezení*.
- Omezení (1.9a) příp. (1.9b) 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 $g_i(\mathbf{x}) \leq 0$ splněno 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*.

Dále uvedeme několik příkladů formulace (a často i řešení) úloh ve tvaru (1.10). Některé z nich byste měli být schopni řešit znalostmi a dovednostmi ze střední školy příp. z analýzy, jde tedy o opakování. Jiné předesílají, co přijde v pozdějších kapitolách.

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

Označme strany obdélníka jako x, y . Obsah obdélníka je xy a jeho obvod (bez strany tvořené řekou) $2x + y$. Řešíme úlohu

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

neboli

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

To je úloha s $n = 2$ proměnnými, $m = 2$ omezeními typu nerovnosti a $l = 1$ omezením typu rovnosti.

I když jde o úlohu s omezeními, bylo by velmi nešikovné snažit se ji řešit formalismem Lagrangeových multiplikátorů či KKT podmínek (pokud je někdo zná). Místo toho z podmínky $2x + y = 100$ vyjádříme $y = 100 - 2x$ (přesněji: rovnice $2x + y = 100$ je ekvivalentní rovnici $y = 100 - 2x$) a dosadíme do původní úlohy. Při eliminaci proměnné y nesmíme zapomenout na nerovnosti: podmínka $y \geq 0$ spolu s $y = 100 - 2x$ implikují $x \leq 50$. Tím dostaneme úlohu s jednou proměnnou

$$\max\{x(100 - 2x) \mid 0 \leq x \leq 50\},$$

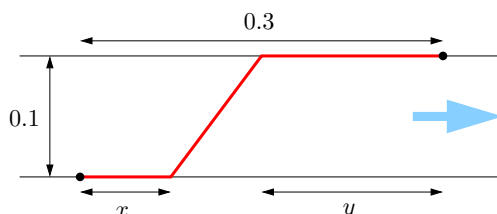
což někdo raději píše jako $\max_{0 \leq x \leq 50} x(100 - 2x)$ nebo $\max_{x \in [0, 50]} x(100 - 2x)$.

Tu snadno vyřešíme metodami analýzy funkcí jedné proměnné. Optimální řešení může být buď ve vnitřním bodě nebo v jednom z krajních bodů intervalu $[0, 50]$. Maximum výrazu

$x(100 - 2x)$ na množině \mathbb{R} snadno najdeme pomocí derivace, nabývá se v bodě $x = 25$. Ten zároveň splňuje podmínku $0 \leq x \leq 50$. Body na krajích intervalu mají menší hodnotu kritéria, tedy nejsou optimální. Dosazením dostaneme $y = 100 - 2x = 50$. ♦

Příklad 1.3. Jste na pravém břehu řeky široké 0.1 km a chcete se dostat ke stanu na levém břehu, který je 0.3 km po proudu od bodu, který je na levém břehu nejbližší vám. Řeka teče pomalu, zanedbatelnou rychlostí. Plavete rychlostí 1 km/h a chodíte rychlostí 3 km/h (běhat odmítáte, protože je vedro). Jaký je nejkratší čas, za který se dokážete dostat ke stanu?

Optimální dráha bude mít tvar jako na obrázku:



Tedy nejdřív jdeme kus x po pravém břehu, pak přeplaveme šikmo do bodu na levém břehu vzdáleném y od stanů, nakonec dojdeme kus y po levém břehu.

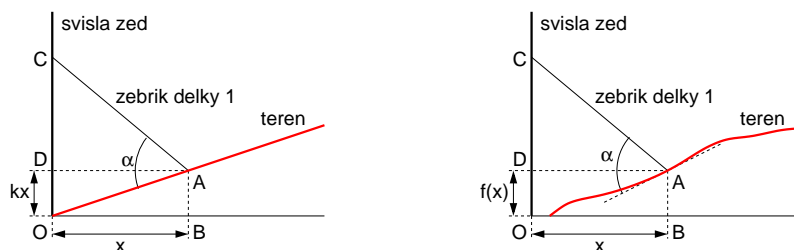
Může mít dráha jiný tvar? Těžko, pokud přijmeme, že dráha, kterou se dostaneme nejrychleji z bodu do bodu za předpokladu, že v každém místě prostoru se pohybujeme stejně rychle, je úsečka spojující tyto dva body.

Celkový čas je $t = \frac{1}{3}x + \sqrt{0.1^2 + (0.3 - x - y)^2} + \frac{1}{3}y$ (hodin). Ten chceme minimalizovat pro proměnné $x, y \in \mathbb{R}$. Všimněte si, že nepotřebujeme omezení $x, y \geq 0$ ani $x + y \leq 0.3$, neboť i bez nich dostaneme přípustné dráhy. Ovšem tyto dráhy očividně nejsou optimální, protože pro $x < 0$ bychom na začátku hloupě kráčeli pryč od stanů, pro $y < 0$ bychom doplávali až za stan, a pro $x + y > 0.3$ bychom zase plavali pryč od stanů. Takže kdybychom tato omezení přidali, úloha by se nezměnila (tj. omezení jsou redundantní).

Máme tedy úlohu se dvěma proměnnými bez omezení. Ovšem proměnné se vyskytují jen v součtu, tedy označíme-li $z = x + y$, je $t = \frac{1}{3}z + \sqrt{0.1^2 + (0.3 - z)^2}$ (samozřejmě vidíme i z obrázku, že x nebo y můžeme uvažovat nulové bez újmy na obecnosti). Pomocí derivace získáme stacionární bod $z = 0.3 - \sqrt{2}/40$ km ≈ 264.6 m. Minimální čas je $t \approx 0.19428$ hodin, tedy asi 11.7 minut. ♦

Příklad 1.4. Zloděj má žebřík délky 1 a potřebuje se dostat přes svislou zeď ze strany, kde terén (v obrázku červeně) stoupá s konstantní (kladnou) směrnici k . Jak daleko od zdi musí zloděj žebřík zapíchnout do země, aby jeho druhý konec dosáhl co nejvýše na zeď? Jaký bude v tom případě úhel mezi žebříkem a terénem?

Nakreslíme situaci (levý obrázek):



Hledáme hodnotu x , která maximalizuje výšku vršku žebříku na zdi $|OC| = kx + \sqrt{1 - x^2}$, za podmínky $x \geq 0$ (protože žebřík nemůžeme zapíchnout za zeď). Tuto podmínku ale můžeme

ignorovat, protože zjevně žádná poloha s $x < 0$ nebude optimální (vršek bude níž než pro jakékoli $x \geq 0$). Tedy opět analýza jedné proměnné: podmínka stacionarity je $k = x/\sqrt{1-x^2}$, z toho $x = 1/\sqrt{1+1/k^2}$. Jedná se o maximum, což můžeme ověřit pomocí druhé derivace (musí být záporná), ale je to i vidět.

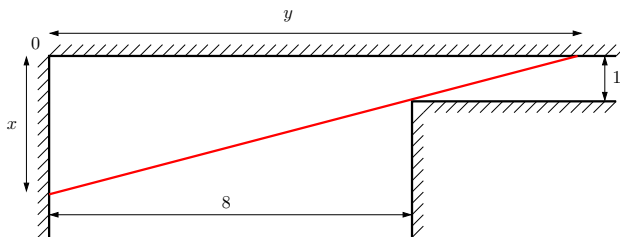
Ukážeme, že pro optimální x bude žebřík kolmý k terénu (tj. $\alpha = \pi/2$). Z podmínky $k = x/\sqrt{1-x^2}$ vidíme, že $|AD|/|CD| = k$, protože $|AD| = x$ a $|CD| = \sqrt{1-x^2}$. Ale z definice směrnice je také $kx/x = |DO|/|AD| = |AB|/|BO| = k$. Tedy pravoúhlé trojúhelníky DCA , BOA a DAO jsou si podobné. Protože součet úhlů v trojúhelníku je π , úhel CAO (tj. α) musí být pravý.

Tento výsledek jsme mohli uhodnout následující úvahou: pokud $\alpha < \frac{\pi}{2}$, pak zmenšením x vršek žebříku C povyleze vzhůru. Podobně pro $\alpha > \frac{\pi}{2}$ a zvětšení x . Když ale $\alpha = \frac{\pi}{2}$, pak zmenšení i zvětšení x způsobí pokles bodu C . Jinými slovy, z obrázku vidíme, že $\alpha = \frac{\pi}{2}$ je lokální maximum.

Úlohu nyní zobecníme tak, že terén nemá tvar přímky $f(x) = kx$, ale obecné neklesající diferencovatelné funkce $f(x)$ (viz obrázek nahoře vpravo). Dosahuje-li žebřík na zdi (bod C) nejvýše, jaký bude úhel (opět označený α) mezi žebříkem a terénem? Minimalizujeme výraz $f(x) + \sqrt{1-x^2}$. Stacionární podmínka je $f'(x) = x/\sqrt{1-x^2}$. To opět znamená, že je žebřík kolmý k terénu (odvození podobné jako minule). ♦

Příklad 1.5. Jdete chodbou o šířce 8 m, která zatáčí do pravého úhlu a zároveň se zužuje na šířku 1 m (viz obrázek). Jaká je největší možná délka rovné tuhé tyče, se kterou lze zatáčkou projít? Požaduje se, abyste tyč drželi stále vodorovně.

Nejdelsí tyč, se kterou projedeme, se určitě bude dotýkat vnitřního rohu chodby (viz obrázek):



Uvažujme pravoúhlý trojúhelník, jehož přepona je tyč (na obrázku červeně) dotýkající se tohoto rohu a odvěsny jsou označeny x a y . Potřebujeme najít délky x a y , pro které bude délka odvěsny minimální (to bude chvíle, kdy bychom už delší tyč neprošli). Trojúhelník má vrcholy $(0, 0)$, $(x, 0)$ a $(0, y)$. Z podobnosti trojúhelníků máme $(x-1)/1 = 8/(y-8)$. Za této podmínky minimalizujeme čtverec délky odvěsny $x^2 + y^2$. Měli bychom přidat ještě podmínky $x, y \geq 0$, ale opět se snadno ukáže, že v optimálním řešení budou splněny.

Z podmínky vyjádříme $y = (8x)/(x-1) = 8/(1-1/x)$ a dosadíme do kritéria, takže minimalizujeme $x^2 + 64/(1-1/x)^2$. Stacionární podmínka je $2x(x^3 - 3x^2 + 3x - 65)/(x-1)^3 = 0$, což je ekvivalentní $2x(x^3 - 3x^2 + 3x - 65) = 0$ za předpokladu $x \neq 1$. Bohužel, to je rovnice 4. stupně. Její jediné reálné řešení je $x = 5$, což se dá uhodnout nebo použít vhodný numerický algoritmus. Z toho $y = 10$. Tedy největší možná délka tyče je $\sqrt{x^2 + y^2} = \sqrt{5^2 + 10^2} = 5\sqrt{5} \approx 11.18034$.

Existuje jiný, jednodušší způsob řešení, kterým se navíc vyhneme rovnici 4. stupně. Délku tyče lze napsat jako $8/\cos \alpha + 1/\sin \alpha$ kde α je úhel mezi tyčí a osou x . Hledáme α , pro které je tato délka minimální. Stacionární podmínka je $2 \sin \alpha = \cos \alpha$, tedy $\tan \alpha = \frac{1}{2}$. Lze ověřit, že jde o minimum. ♦

Příklad 1.6. Na zahradě vám rostou melouny. Dnes je jich tam 200 kg. Každý den melouny narostou o 5 kg, ale cena 1 kg melounů na trhu klesne o 10 haléřů. Pokud dnešní cena 1 kg melounů na trhu je 9 Kč, jak dlouho máte čekat se sklizní, abyste dosáhli co největšího zisku? Předpokládejte, že melouny sklídíte a prodáte tentýž den.

Cena na trhu za t dnů je $(200 + 5t)(9 - t/10) = -t^2/2 + 25t + 1800$. Chceme najít takové t , pro které tato funkce bude maximální. Porovnáním derivace s nulou dostaneme stacionární bod $t = 25$, který je maximem.

Ovšem pozor: správně jsme měli požadovat, aby t bylo nezáporné a celočíselné, tedy $t \geq 0$ a $t \in \mathbb{Z}$. V našem případě jsme měli štěstí a optimální řešení t tyto podmínky splňuje. I kdyby je nesplňovalo, snadno bychom z něj spočetli optimální řešení, které by je splňovalo (jak?).

Prakticky zaměřený čtenář namítne, že žádný zelinář by to takto nedělal. Hmotnost melounů ani jejich výkupní cena nejsou lineární funkce času a skutečné funkce není snadné odhadnout. I kdyby je zelinář znal, jeho kritériem není jen velikost zisku z melounů, protože např. pěstuje i spoustu jiných věcí a v den optimální sklizně melounů zrovna musí okopávat okurky.

Většina slovních úloh v těchto skriptech bude takto nerealisticky zjednodušená (půjde o tzv. *toy problems*). Skutečné úlohy z optimalizační praxe jsou obvykle příliš složité na to, abychom je v tomto kursu dokázali uvést a vysvětlit. Tato složitost je ovšem často jen kvantitativního rázu (více proměnných a více omezení), *typ* úlohy je stejný jako u našich zjednodušených úloh. ♦

Příklad 1.7. Řešme tuto lineární soustavu třech rovnic o dvou neznámých:

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

Tato soustava nemá řešení (je přeuročená). Chceme najít alespoň přibližné řešení.

Termín ‘přibližné řešení’ není jednoznačný, různí lidé jím mohou myslet různé věci. Nejčastěji užívaná formulace je minimalizovat součet čtverců⁴ zbytků (residuí) jednotlivých rovnic. Tedy minimalizujeme funkci

$$f(x, y) = (x + 2y - 6)^2 + (-x + y - 3)^2 + (x + y - 4)^2$$

na množině \mathbb{R}^2 . Je jasné, že $f(x, y) > 0$ pro všechna $x, y \in \mathbb{R}$, kdyby totiž $f(x, y) = 0$ pak by x, y bylo řešení soustavy. Minimum funkce f najdeme snadno, protože f je kvadratická funkce dvou proměnných. Nutná (a zde i postačující) podmínka na minimum je nulovost parciálních derivací.

Jiný způsob, jak formalizovat termín ‘přibližné řešení’, je minimalizovat funkci

$$f(x, y) = |x + 2y - 6| + |-x + y - 3| + |x + y - 4|.$$

Zde už nám derivace nepomohou. Uvidíme, že úloha je konvexní a dá se převést na lineární programování. ♦

Příklad 1.8. Víme, že reálná veličina y závisí na reálné veličině x kvadraticky, tj. $y = ax^2 + bx + c$. Neznáme ovšem koeficienty a, b, c . Naměřili jsme hodnoty veličiny y pro m hodnot veličiny x , tj. máme dvojice $(x_i, y_i)_{i=1}^m$. Jak najdeme koeficienty a, b, c ?

⁴‘Čtverec’ čísla znamená jeho druhou mocninu. Tento historický termín je jednak hezký a jednak se používá tak často, že ho budeme používat i my.

To je úloha na lineární regresi. Lze ji formulovat ve smyslu nejmenších čtverců jako minimalizaci funkce

$$f(a, b, c) = \sum_{i=1}^m (ax_i^2 + bx_i + c - y_i)^2$$

přes proměnné $a, b, c \in \mathbb{R}$. Tato formulace má statistické odůvodnění (z principu maximální věrohodnosti) za předpokladu, že čísla x_1, \dots, x_n měříme přesně a nepřesnosti měření čísel y_1, \dots, y_m jsou i.i.d. normální s nulovou střední hodnotou a stejnou variancí. Minimum kvadratické funkce f opět najdeme porovnáním parciálních derivací s nulou. ♦

Příklad 1.9. Jsou dány body $\mathbf{a}_1 \neq \mathbf{b}_1$ a $\mathbf{a}_2 \neq \mathbf{b}_2$ v prostoru \mathbb{R}^n . Najděte vzdálenost přímky procházející body $\mathbf{a}_1, \mathbf{b}_1$ od přímky procházející body $\mathbf{a}_2, \mathbf{b}_2$ (tj. vzdálenost mimoběžek, neboli délka přímky mimoběžek).

Víme, že i -tá (kde $i = 1, 2$) přímka je množina $\{\mathbf{a}_i + \alpha(\mathbf{b}_i - \mathbf{a}_i) \mid \alpha \in \mathbb{R}\}$. Hledáme bod na první přímce a bod na druhé přímce tak, aby jejich vzdálenost byla co nejmenší. Tedy minimalizujeme funkci

$$f(\alpha_1, \alpha_2) = \|\mathbf{a}_1 + \alpha_1(\mathbf{b}_1 - \mathbf{a}_1) - \mathbf{a}_2 - \alpha_2(\mathbf{b}_2 - \mathbf{a}_2)\|^2, \quad (1.14)$$

kde

$$\|\mathbf{x}\| = (x_1^2 + \dots + x_n^2)^{1/2}$$

označuje eukleidovskou normu vektoru $\mathbf{x} \in \mathbb{R}^n$. Funkce (1.14) je opět kvadratická funkce dvou proměnných. ♦

Příklad 1.10. Pán⁵ 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 hranolků se spotřebuje 1.5 kg brambor a 0.2 kg oleje. Je nakoupeno 100 kg brambor a 16 kg oleje. Kolik má pán vyrobit lupínků a kolik hranolků, aby co nejvíce utržil? Přitom nepočítáme cenu surovin a předpokládáme, že všechny výrobky se prodají a nevyužitá surovina se po pracovní době vyhodí.

Tuto úlohu lze formalizovat takto:

$$\begin{aligned} \max \quad & 120l + 76h \\ \text{za podmíněk} \quad & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l, h \geq 0 \end{aligned}$$

Optimální řešení je $l = 20$ kg lupínků a $h = 40$ kg hranolků. Podotkněme, že v tomto příkladě jsou omezení $x, y \geq 0$ zbytečná – kdybychom je vynechali, optimální řešení by se nezměnilo (to je vidět z toho, že ani jedna z podmínek $x, y \geq 0$ není v optimu aktivní). Ovšem kdyby druhá surovina a výrobků bylo více, podmínky by zbytečné nebyly. ♦

Příklad 1.11. Najděte nejmenší n -rozměrnou kouli obsahující dané body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$.

⁵Úloha převzata z elektronických skript *Petr Hliněný: Optimalizační úlohy. Kat. informatiky VSB-TU Ostrava, 2006.*

Koule se středem \mathbf{x} a poloměrem r je množina $\{\mathbf{a} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\| \leq r\}$. Úlohu lze tedy napsat jako

$$\begin{aligned} \min \quad & r \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\| \leq r, \quad i = 1, \dots, m \\ & \mathbf{x} \in \mathbb{R}^n \\ & r \in \mathbb{R} \end{aligned} \tag{1.15}$$

Mohla by tam být ještě podmínka $r \geq 0$, ta je ale zbytečná, protože norma je nezáporná. Úloha (1.15) je konvexní a patří do třídy programování na kuželu druhého řádu (SOCP), což si ovšem řekneme až mnohem později.

Protože minimalizaci přes proměnnou r lze udělat explicitně, úloha (1.15) je ekvivalentní minimalizaci funkce

$$f(\mathbf{x}) = \max_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\| \tag{1.16}$$

na množině \mathbb{R}^n (tj. bez omezení). To lze napsat také jako

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|. \tag{1.17}$$

Umocněním podmínek na druhou a substitucí $r^2 = y$ můžeme úlohu napsat ještě v jiném ekvivalentním tvaru

$$\begin{aligned} \min \quad & y \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\|^2 \leq y, \quad i = 1, \dots, m \\ & \mathbf{x} \in \mathbb{R}^n \\ & y \in \mathbb{R} \end{aligned} \tag{1.18}$$

neboli (po eliminaci y)

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|^2. \tag{1.19}$$

Úloha (1.18) je konvexní úloha kvadratického programování (QCQP), o které si také později řekneme.

Na konvexní QCQP a SOCP existují numerické metody, které bychom mohli mechanicky použít. Ovšem na tuto úlohu byly vymyšleny i speciální algoritmy (hlavně pro případy $n = 2$ a $n = 3$), které jsou rychlejší.

Příklad 1.12. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a hledáme vektor $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, který minimalizuje (bez omezujících podmínek) funkci

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|^2. \tag{1.20}$$

Označíme-li $\mathbf{a}_i = (a_{i1}, \dots, a_{in})$, je $\sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|^2 = \sum_{i=1}^m \sum_{j=1}^n (x_j - a_{ij})^2$, tedy účelová funkce je součtem n funkcí, z nichž každá závisí jen na jedné souřadnici x_j . Minimum funkce f lze tedy najít tak, že najdeme minimum každé funkce zvlášť (úloha se nám tak ‘rozpadla’ na n nezávislých optimalizačních úloh). Jak snadno spočtete (a jistě jste dávno uhodli), minimum se nabývá v *těžišti* $(\mathbf{a}_1 + \dots + \mathbf{a}_m)/m$ daných bodů. \blacklozen

Příklad 1.13. Jsou dána čísla $a_1, \dots, a_m \in \mathbb{R}$. Najděte minimum funkce

$$f(x) = \sum_{i=1}^m |x - a_i|. \quad (1.21)$$

Seřadíme čísla a_1, \dots, a_m vzestupně, tedy předpokládáme $a_1 \leq a_2 \leq \dots \leq a_{m-1} \leq a_m$. Funkce f je po částech lineární (přesněji: je afinní), je diferencovatelná všude kromě bodů a_i . Derivace je konstantní pro každý interval (a_{i-1}, a_i) . Najděte hodnoty těchto derivací. K tomu si nejprve ujasněte, jak vypadá funkce $|x - a_i|$ pro jediné i a jaká je její derivace (derivaci můžete napsat pomocí funkce signum).

Z derivací na intervalech usud'te, kde je funkce f klesající, kde rostoucí a kde konstantní. Nyní je jasné, kde f nabývá minima: pro m liché je to v bodě $a_{(m+1)/2}$, pro m sudé na intervalu $[a_{m/2}, a_{m/2+1}]$. Argument minima se označuje jako *medián* čísel a_1, \dots, a_m (pro sudé m se konvencí bere číslo $\frac{1}{2}(a_{m/2}, a_{m/2+1})$). ♦

Příklad 1.14. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Najděte minimum funkce

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|. \quad (1.22)$$

Pro $n = 1$ se funkce (1.22) redukuje na (1.21). Pro $n \geq 2$ je řešení úlohy známo jako *geometrický medián*. Ovšem na rozdíl od obyčejného mediánu se minimum obecně nenabývá v žádném z bodů $\mathbf{a}_1, \dots, \mathbf{a}_m$. Neexistuje algoritmus, který by pro $n \geq 2$ našel minimum funkce (1.22) v konečném počtu kroků. Dobrá zpráva ale je, že úloha je konvexní (patří do třídy SOCP, o které jsme se už zmínili).

Pro případ $n = 2$ má úloha jednoduchý mechanický model⁶. Do vodorovného prkna vyvrátíme 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 $f(\mathbf{x})$. ♦

Příklad 1.15. Mějme m bodů $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ v n -rozměrném prostoru. Úkolem je rozmístit dalších k bodů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ tak, aby průměrná vzdálenost bodu \mathbf{a}_i k nejbližšímu bodu \mathbf{x}_j byla co nejmenší. Tedy minimalizujeme účelovou funkci

$$f(\mathbf{x}_1, \dots, \mathbf{x}_k) = \sum_{i=1}^m \min_{j=1}^k \|\mathbf{a}_i - \mathbf{x}_j\| \quad (1.23)$$

pro neznámé vektory $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$. Úloha je známá jako *shlukování* (angl. *clustering*). 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.

Není znám (a pravděpodobně nikdy nebude) algoritmus, který by pro libovolný vstup (tedy libovolné $n, m, k \in \mathbb{N}$ a $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$) našel globální minimum funkce (1.23) za prakticky přijatelný čas. Lze totiž dokázat, že úloha je tzv. NP-těžká⁷. V praktické situaci často použijeme algoritmus, který najde pouze přibližné (typicky lokální) optimum, např. *k-means*. ♦

⁶Toto 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.

⁷Tento pojem patří do *teorie složitosti algoritmů*, kterou jste ještě nebrali. Zde jen řekneme, že pro NP-těžkou úlohu jen malá naděje, že bude někdy nalezen algoritmus, který by úlohu shlukování řešil v polynomiálním čase. Algoritmus řeší úlohu v *polynomiálním čase*, jestliže existuje polynom p takový, že pro každý vstup najde řešení v čase menším než $p(L)$, kde L je počet bitů potřebných k zápisu vstupu.

Příklad 1.16. Necht' (V, E) je ohodnocený neorientovaný graf, tj. V je konečná množina, $E \subseteq \binom{V}{2}$ je množina dvouprvkových podmnožin V , a je dáno zobrazení $c: E \rightarrow \mathbb{R}$ (tedy každá hrana $\{i, j\} \in E$ má přiřazené číslo c_{ij}). Úloha na *maximální řez* (*maximum cut*)⁸ v grafu zní

$$\begin{aligned} & \max \sum_{\{i,j\} \in E} c_{ij} x_i x_j \\ & \text{za podmínek } \begin{cases} x_i^2 = 1, & i \in V \\ x_i \in \mathbb{R}, & i \in V \end{cases} \end{aligned} \quad (1.24)$$

Účelová funkce je kvadratická (dokonce bilineární) a máme $|V|$ kvadratických omezení typu rovnosti. Každé omezení $x_i^2 = 1$ je ekvivalentní $x_i \in \{-1, 1\}$, tedy množina přípustných řešení má konečný počet ($2^{|V|}$) prvků a jedná se tedy o kombinatorickou úlohu. Všimněte si že, přísně vzato, omezení nemůžeme psát jako $x_i \in \{-1, 1\}$, protože to není ve tvaru (1.10). Jedná se o klasickou NP-těžkou úlohu, proto už pro některé poměrně malé úlohy je prakticky nemožné najít globální optimum. ♦

1.4 Cvičení

1.1. Vyřešte následující úlohy, přičemž slovní úlohy nejdříve formulujte ve tvaru (1.10). Stačí vám k tomu papír, tužka, zdravý rozum a analýza funkcí jedné proměnné. Všimněte si, že některé úlohy lze převést na hledání extrémů funkce jedné proměnné na intervalu, což umíte z analýzy funkcí jedné proměnné.

- $\min\{x^2 + y^2 \mid x \geq 0, xy \geq 1\}$
- $\min\{(x-2)^2 + (y-\frac{1}{2})^2 \mid x^2 \leq 1, y^2 \leq 1\}$
- $\min\{x \mid x \in \mathbb{R}, x \geq a_i \forall i = 1, \dots, n\}$ pro daná $a_1, \dots, a_n \in \mathbb{R}$
- Máte vyrobit papírovou krabici (včetně víka) o objemu 72 litrů, jejíž délka je dvojnásobek její šířky. Jaké budou její rozměry, má-li se na ní spotřebovat co nejméně papíru? Tloušťka stěn je zanedbatelná.
- Jaké má rozměry válec s jednotkovým objemem a nejmenším povrchem?
- Najděte rozměry pultitru, na jehož výrobu je třeba co nejméně skla. Tloušťka stěn je zanedbatelná.
- Najděte obsah největšího obdélníka vepsaného do kružnice s poloměrem 1.
- Obdélník v rovině má jeden roh v počátku a druhý na křivce $y = x^2 + x^{-2}$, přičemž jeho strany jsou rovnoběžné se souřadnicovými osami. Pro jaké x bude jeho obsah minimální? Může být jeho obsah libovolně veliký?
- Najděte bod v rovině na parabole s rovnicí $y = x^2$ nejbližší bodu $(3, 0)$.
- Hektarová oblast obdélníkového tvaru se má obehnat ze tří stran živým plotem, který stojí 1000 korun na metr, a ze zbývajících stran obyčejným plotem, který stojí 500 korun na metr. Jaké budou rozměry oblasti při nejmenší ceně plotu?
- x, y jsou čísla v intervalu $[1, 5]$ taková, že jejich součet je 6. Najděte tato čísla tak, aby xy^2 bylo (a) co nejmenší a (b) co největší.

⁸Úloha byla intenzivně studována nejen v kombinatorické optimalizaci, ale také ve statistické fyzice pod názvem hledání *minimální energie Isingova modelu*.

- l) 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 .
- m) *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.
- n) Normální rozdělení se střední hodnotou μ a směrodatnou odchylkou σ má hustotu pravděpodobnosti $p_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. Chceme odhadnout parametry μ a σ z i.i.d. vzorku x_1, \dots, x_n z rozdělení na základě principu maximální věrohodnosti, tedy chceme maximalizovat věrohodnost $\prod_{i=1}^n p_{\mu,\sigma}(x_i)$.
- 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 $\operatorname{argmin}_{x \in X} f(x) = \operatorname{argmin}_{x \in X} g(f(x))$.
- 1.3. Necht' X je libovolná množina, $Y \subseteq X$, a $f: X \rightarrow \mathbb{R}$. Najděte co nejobecnější podmínku, za které platí $\operatorname{argmin}_{x \in Y} f(x) = Y \cap \operatorname{argmin}_{x \in X} f(x)$.
- 1.4. Dokažte, že $\operatorname{argmin}_{x \in X} f(x) = \operatorname{argmin}_{x \in X} g(f(x))$ jestliže $g: \mathbb{R} \rightarrow \mathbb{R}$ je rostoucí funkce.

Nápověda a řešení

- 1.1.a) Vyřešíme úvahou s pomocí obrázku. Množina přípustných řešení je kladná větev hyperboly $xy = 1$ a oblast nad ní. Účelová funkce $f(x, y) = x^2 + y^2$ je čtverec (= druhá mocnina) vzdálenosti bodu (x, y) od počátku, její vrstevnice jsou kružnice se středem v počátku $(0, 0)$. Protože odmocnina je rostoucí funkce, úloha (přesněji, její množina argumentů minima) by se nezměnila (viz Cvičení 1.2), kdybychom účelovou funkce změnili na $\sqrt{x^2 + y^2}$, což je vzdálenost od počátku. Optimální hodnota se proto nabývá v bodě hyperboly nejbliž počátku, tedy v bodě $(x, y) = (1, 1)$.
- 1.1.b) Vyřešíme opět obrázkem. Protože $x^2 \leq 1$ je ekvivalentní $-1 \leq x \leq 1$, množina přípustných řešení je čtverec (i s vnitřkem) se středem v počátku a stranou délky 2. Účelová funkce je čtverec vzdálenosti od bodu $(2, \frac{1}{2})$. Minimum se nabývá v bodě čtverce nejbliž bodu $(2, \frac{1}{2})$, tj. v bodě $(1, \frac{1}{2})$.
- 1.1.c) Optimální řešení je nejmenší číslo x , které není menší než žádné z čísel a_1, \dots, a_n . Očividně, takové číslo je $x = \max_i a_i$.
- 1.1.d) Krabice je kvádr se stranami $x, 2x, y$ (v decimetrech). Minimalizujeme jeho povrch $6x + 2y$ za podmíněk $2x^2y = 72$ a $x, y \geq 0$. Z této úlohy eliminujeme proměnnou y . První podmínka je ekvivalentní $y = 36/x^2$. Všimneme si, že podmínka $y = 36/x^2 \geq 0$ je automaticky splněna pro $x > 0$. Tedy minimalizujeme $6x + 72/x^2$ za podmínky $x > 0$. To spadá do analýzy funkcí jedné proměnné: položení derivace rovné nule dá stacionární bod $x = \sqrt[3]{24} = 2\sqrt[3]{3}$, což splňuje $x \geq 0$. Mohli bychom ověřit, např. pomocí druhé derivace nebo úvahou, že je to globální minimum na intervalu $[0, \infty)$.
- 1.1.h) Minimalizujeme obsah xy obdélníka za podmínky $y = x^2 + x^{-2}$ přes proměnné $x, y \in \mathbb{R}$, $x \neq 0$ (poslední podmínka je kvůli x^{-2}). Dosazením za y převedeme na minimalizaci funkce jedné proměnné $x(x^2 + x^{-2}) = x^3 + 1/x$ (bez omezení, tedy na intervalu $(-\infty, \infty)$).
- 1.1.i) Minimalizujeme čtverec vzdálenosti $(x - 3)^2 + y^2$ za podmínky $y = x^2$ přes proměnné $x, y \in \mathbb{R}$. Dosazením za y převedeme na minimalizaci funkce $(x - 3)^2 + x^4$ bez omezení.

- 1.1.k) Hledáme extrém funkce xy^2 za podmínek $x + y = 6$ a $x, y \in [1, 5]$. Tuto úlohu převedeme na úlohu s jedinou proměnnou eliminací např. y . Z $x + y = 6$ vyjádříme $y = 6 - x$. Všimneme si (ověřte podrobně!), že podmínka $6 - x \in [1, 5]$ je ekvivalentní $x \in [1, 5]$, což už ale máme. Tedy minimalizujeme $x(6 - x)^2$ na intervalu $x \in [1, 5]$.
- 1.1.l) Minimalizujeme $x_1 + \dots + x_n$ za podmínek $x_1 \dots x_n > 0$ a $x_1, \dots, x_n \in \{-1, 1\}$. To není úloha ve tvaru (1.10), protože omezení $x_i \in \{-1, 1\}$ nejsou v (1.10) dovolena. To ovšem snadno spravíme, protože je můžeme nahradit ekvivalentními omezeními $x_i^2 = 1$, které už tvar (1.10) dovoluje. Stejně ale jde o netypickou úlohu spojité optimalizace, protože každá proměnná může nabývat jen dvou hodnot a tedy množina přípustných řešení má konečný počet prvků (jedná se tedy spíše o úlohu kombinatorické optimalizace).
- Úlohu vyřešíme jednoduchou úvahou. Aby $x_1 \dots x_n > 0$, zápornou hodnotu smí mít sudý počet proměnných. Chceme-li minimalizovat $x_1 + \dots + x_n$, pro sudé n tedy nastavíme všechny (pro sudé n) příp. všechny kromě jedné (pro liché n) proměnné na -1 . Optimální hodnota bude tedy $-n$ (pro n sudé) a $1 - n$ (pro n liché).
- 1.1.n) Tedy chceme maximalizovat $\prod_{i=1}^n p_{\mu, \sigma}(x_i)$ přes $\mu \in \mathbb{R}$ a $\sigma > 0$. Je šikovné místo maximalizace této funkce minimalizovat její záporný logaritmus.
- 1.2. Úkolem je dokázat, že pro každé $x^* \in X$ platí následující ekvivalence: pro každé $x \in X$ platí $f(x^*) \leq f(x)$ právě tehdy, když pro každé $x \in X$ platí $g(f(x^*)) \leq g(f(x))$. To je ale zjevné, protože pro každou rostoucí funkci $g: \mathbb{R} \rightarrow \mathbb{R}$ a každé $a, b \in \mathbb{R}$ platí $a \leq b \Leftrightarrow g(a) \leq g(b)$.

Část I

Použití lineární algebry v optimalizaci

Kapitola 2

Maticová algebra

Cílem této kapitoly je zopakovat a prohloubit si schopnost manipulovat s výrazy obsahujícími matice a vektory, aniž bychom zatím používali abstraktnější pojmy z lineární algebry jako např. lineární nezávislost.

Reálná **matice** rozměru $m \times n$, kde $m, n \in \mathbb{N}$, je tabulka reálných čísel s m řádky a n sloupci¹,

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

kde a_{ij} jsou **prvky** matice². Množinu všech reálných matic rozměru $m \times n$ značíme $\mathbb{R}^{m \times n}$.

Používají se 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 $\mathbf{0}_{m,n}$ (pokud jsou rozměry m, n jasné z kontextu, pak pouze $\mathbf{0}$).
- **Jednotková matice** je čtvercová diagonální, jejíž diagonální prvky jsou jedničky. Značíme ji \mathbf{I}_n (pokud je rozměr n jasná z kontextu, pak pouze \mathbf{I}). Prvky jednotkové matice můžeme označit symbolem 'Kroneckerovo delta',

$$\delta_{ij} = \begin{cases} 0 & \text{když } i \neq j, \\ 1 & \text{když } i = j. \end{cases} \quad (2.1)$$

- **Horní [dolní] trojúhelníková matice** má $a_{ij} = 0$ pro všechna $i > j$ [$i < j$]. Všimněte si, že horní/dolní trojúhelníková matice nemusí být čtvercová.
- **Permutační matice** je čtvercová matice, která má v každém sloupci právě jednu jedničku, v každém řádku právě jednu jedničku, a ostatní jsou nuly.

¹Formálněji můžeme matici definovat jako zobrazení $\{1, \dots, m\} \times \{1, \dots, n\} \rightarrow \mathbb{R}$, podobně jako vektor $(x_1, \dots, x_n) \in \mathbb{R}^n$ lze považovat za zobrazení $\{1, \dots, n\} \rightarrow \mathbb{R}$.

²Někdo značí prvky matice \mathbf{A} jako A_{ij} , což má své výhody, my se ale budeme držet častějšího a_{ij} .

2.1 Operace s maticemi

2.1.1 Sčítání a násobení skalárem

Matice lze násobit reálným číslem a sčítat po prvcích:

- Součin čísla $\alpha \in \mathbb{R}$ 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 $\frac{1}{\alpha} \mathbf{A}$ můžeme psát krátce jako $\frac{\mathbf{A}}{\alpha}$ nebo \mathbf{A}/α . Součin $(-1)\mathbf{A}$ píšeme krátce jako $-\mathbf{A}$.

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

Pro pevná $m, n \in \mathbb{N}$ tvoří množina $\mathbb{R}^{m \times n}$ vybavená těmito dvěma operacemi lineární prostor nad tělesem \mathbb{R} . Prvky tohoto tělesa, tedy čísla $\alpha \in \mathbb{R}$, se nazývají **skaláry**. Zdůrazněme, že slovo *skalár* nemá význam samostatně, ale vždy jen v kontextu nějakého lineárního prostoru.

2.1.2 Maticový součin

Maticový součin matic $\mathbf{A} \in \mathbb{R}^{m \times p}$ a $\mathbf{B} \in \mathbb{R}^{p \times n}$ je matice $\mathbf{AB} = \mathbf{C} \in \mathbb{R}^{m \times n}$ s prvky

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \quad (2.2)$$

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

- $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ (asociativita)
- $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ a $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ (distributivita se sčítáním)
- $\mathbf{AI} = \mathbf{A} = \mathbf{IA}$ (neutralita jednotkové matice)
- $(\alpha \mathbf{A})\mathbf{B} = \mathbf{A}(\alpha \mathbf{B}) = \alpha(\mathbf{AB})$ (asociativita s násobením skaláry)

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

Zdůrazněme, že skalár nelze považovat za ‘matici’ rozměru 1×1 , i když to k tomu svádí. Výraz $\alpha \mathbf{A}$ není maticový součin, už proto že vnitřní rozměr matic by byl obecně různý. Násobení matice skalárem je zkrátka jiná operace než maticový součin.

Pro čtvercovou matici \mathbf{A} se maticový součin

$$\mathbf{A}^k = \underbrace{\mathbf{AA} \cdots \mathbf{A}}_{k\text{-krát}} \quad (2.3)$$

nazývá k -tá **mocnina** matice. Definujeme $\mathbf{A}^0 = \mathbf{I}$. Odlišujte od kartézského součinu množiny A se sebou k -krát, který také značíme A^k (viz §1.1.1).

2.1.3 Transpozice

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 $a_{ii} = 0$).

2.1.4 Inverze

Když pro matice $\mathbf{A} \in \mathbb{R}^{n \times m}$ a $\mathbf{B} \in \mathbb{R}^{m \times n}$ platí

$$\mathbf{AB} = \mathbf{I}_n, \quad (2.4)$$

matice \mathbf{B} se nazývá **pravá inverze** matice \mathbf{A} a matice \mathbf{A} se nazývá **levá inverze** matice \mathbf{B} . Matice může mít obě tyto inverze, nebo jen jednu z nich, nebo nemusí mít ani jednu. Když má matice pravou nebo levou inverzi, tato nemusí být jediná. Následující tvrzení budeme schopni dokázat³ až v §3.2.3, zatím si je pamatujte:

- Pravá inverze matice existuje, právě když má matice lineárně nezávislé řádky. Speciálně, je-li matice úzká, nemůže mít pravou inverzi.
- Levá inverze matice existuje, právě když má matice lineárně nezávislé sloupce. Speciálně, je-li matice široká, nemůže mít levou inverzi.

Má-li matice pravou i levou inverzi, pak se obě inverze rovnají a jsou jediné. Opravdu, z $\mathbf{AB} = \mathbf{I}$ a $\mathbf{CA} = \mathbf{I}$ plyne $\mathbf{C} = \mathbf{CAB} = \mathbf{B}$. Pak mluvíme pouze o **inverzi** matice a značíme ji \mathbf{A}^{-1} . Z tvrzení výše navíc plyne, že \mathbf{A} i \mathbf{A}^{-1} jsou čtvercové. Matice, která má inverzi, se nazývá **invertovatelná** nebo **regulární**. Čtvercová matice, která nemá inverzi, se nazývá **singulární**. Vlastnosti inverze:

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

2.1.5 Determinant

Determinant čtvercové matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ lze definovat vzorcem

$$\det \mathbf{A} = \sum_{\sigma} \operatorname{sgn} \sigma \prod_{i=1}^n a_{i\sigma(i)}, \quad (2.5)$$

kde sčítáme přes všechny permutace σ prvků $1, \dots, n$ (tj. bijekce $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$), přičemž $\operatorname{sgn} \sigma$ 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 $\mathbf{B} = \mathbf{A}^{-1}$)

³Je zajímavé, že pouze z definice (2.4) (tedy bez použití pojmů jako lineární nezávislost, hodnost nebo lineární zobrazení) nejspíš nelze dokázat, že úzká matice nemůže mít pravou inverzi. Schválně, zkuste najít důkaz!

- $\det \mathbf{A}^T = \det \mathbf{A}$
- $\det \mathbf{A} = 0$ 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 libovolných dvou sloupců změni znaménko determinantu.

Lze dokázat, že první a poslední vlastnost determinant jednoznačně definují, tj. není žádná jiná funkce $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ s těmito dvěma vlastnostmi.

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} . (Vyhovuje i pro $f(\mathbf{X}) = \mathbf{I}$.)

2.1.6 Stopa

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

$$\operatorname{tr} \mathbf{A} = a_{11} + \cdots + a_{nn}. \quad (2.6)$$

Vlastnosti (dokažte!):

- $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B}$
- $\operatorname{tr}(\alpha \mathbf{A}) = \alpha \operatorname{tr} \mathbf{A}$
- $\operatorname{tr}(\mathbf{A}^T) = \operatorname{tr} \mathbf{A}$
- $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$ 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. $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{CAB})$, protože $\operatorname{tr}(\mathbf{DC}) = \operatorname{tr}(\mathbf{CD})$ kde $\mathbf{D} = \mathbf{AB}$. Podobně např. $\operatorname{tr}(\mathbf{ABCD}) = \operatorname{tr}(\mathbf{CDAB})$. Ale neplatí např. $\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{CBA})$.

2.1.7 Standardní skalární součin a norma

Standardní skalární součin matic $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ se definuje jako

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}. \quad (2.7)$$

Standardní skalární součin lze počítat pomocí stopy a naopak:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{A}^T \mathbf{B}) \quad (2.8)$$

$$\operatorname{tr} \mathbf{A} = \langle \mathbf{I}, \mathbf{A} \rangle \quad (2.9)$$

(pro důkaz (2.8) si napište diagonální prvky matice $\mathbf{A}^T \mathbf{B}$).

Z definice (2.7) snadno dokážeme následující vlastnosti skalárního součinu:

- $\langle \mathbf{A}, \mathbf{B} \rangle = \langle \mathbf{B}, \mathbf{A} \rangle$ (komutativita)
- $\langle \mathbf{A}, \mathbf{B} \rangle = \langle \mathbf{A}^T, \mathbf{B}^T \rangle$
- $\langle \alpha \mathbf{A}, \mathbf{B} \rangle = \alpha \langle \mathbf{A}, \mathbf{B} \rangle = \langle \mathbf{A}, \alpha \mathbf{B} \rangle$
- $\langle \mathbf{A} + \mathbf{B}, \mathbf{C} \rangle = \langle \mathbf{A}, \mathbf{C} \rangle + \langle \mathbf{B}, \mathbf{C} \rangle$ a $\langle \mathbf{A}, \mathbf{B} + \mathbf{C} \rangle = \langle \mathbf{A}, \mathbf{B} \rangle + \langle \mathbf{A}, \mathbf{C} \rangle$

Méně očividné je chování skalárního součinu vzhledem k maticovému součinu⁴:

$$\langle \mathbf{AB}, \mathbf{C} \rangle = \langle \mathbf{B}, \mathbf{A}^T \mathbf{C} \rangle = \langle \mathbf{A}, \mathbf{CB}^T \rangle. \quad (2.10)$$

První identitu dokážeme např. takto:

$$\langle \mathbf{AB}, \mathbf{C} \rangle = \langle \mathbf{C}, \mathbf{AB} \rangle = \text{tr}(\mathbf{C}^T \mathbf{AB}) = \text{tr}((\mathbf{A}^T \mathbf{C})^T \mathbf{B}) = \langle \mathbf{A}^T \mathbf{C}, \mathbf{B} \rangle = \langle \mathbf{B}, \mathbf{A}^T \mathbf{C} \rangle.$$

Druhou identitu dokážeme podobně (proved'te!).

Skalární součin (2.7) indukuje **normu matice** $\mathbf{A} \in \mathbb{R}^{m \times n}$ jako

$$\|\mathbf{A}\| = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \sqrt{\text{tr}(\mathbf{A}^T \mathbf{A})} = \sqrt{\text{tr}(\mathbf{AA}^T)} = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{1/2}. \quad (2.11)$$

Existují i jiné maticové normy (které neuvádíme). Maticová norma (2.11) se nazývá **Frobeniova norma** a pokud ji chceme odlišit od jiných norm, značíme ji $\|\cdot\|_F$.

2.2 Matice s jedním sloupcem nebo jedním řádkem

Matici s jediným sloupcem (tedy prvku $\mathbb{R}^{n \times 1}$) se také říká **sloupcový vektor** a matici s jediným řádkem (prvku $\mathbb{R}^{1 \times n}$) **řádkový vektor**. Lineární prostor $\mathbb{R}^{n \times 1}$ matic s jediným sloupcem je ‘skoro stejný’ (je s ním *isomorfní*) jako lineární prostor \mathbb{R}^n uspořádaných n -tic (x_1, \dots, x_n) . Proto je zvykem tyto dva prostory ztotožnit a bez upozornění přecházet mezi oběma významy. Prvkům

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

tohoto prostoru budeme říkat krátce **vektory**. Slovem *vektor* (bez přívlastku) tedy rozumíme *sloupcový vektor* (tj. matici s jedním sloupcem) nebo uspořádanou n -tici čísel. Zopakujme, že uspořádané n -tice (prvky kartézského součinu) značíme *kulatými závorkami* (viz §1.1.3), kdežto matice značíme *hranatými závorkami*.

Totéž bychom samozřejmě mohli udělat s řádky (tj. ztotožnit prostory \mathbb{R}^n a $\mathbb{R}^{1 \times n}$), ale převládající konvence je dělat to se sloupci⁵. Chceme-li tedy napsat řádkový vektor, můžeme napsat $\mathbf{x} \in \mathbb{R}^{1 \times n}$, ale to se většinou nedělá, neboť malá tučná písmena obvykle označují prvky \mathbb{R}^n . Řádkový vektor proto raději píšeme jako \mathbf{x}^T kde $\mathbf{x} \in \mathbb{R}^n$, z čehož plyne $\mathbf{x}^T \in \mathbb{R}^{1 \times n}$.

Popsané ztotožnění je důležité, protože nyní se k vektorům $\mathbf{x} \in \mathbb{R}^n$ můžeme chovat jako k maticím, tj. můžeme je maticově násobit jinými maticemi, transponovat, apod. Uveďme důležité případy, kdy se v maticovém součinu vyskytují vektory:

- Pro $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{x} \in \mathbb{R}^n$ je výraz \mathbf{Ax} maticový součin matice $m \times n$ a matice $n \times 1$, což je (sloupcový) vektor délky m . Je to vlastně lineární kombinace sloupců matice \mathbf{A} s koeficienty \mathbf{x} (viz §3.2).

⁴V obecném lineárním prostoru se skalárním součinem vlastnost $\langle f(x), y \rangle = \langle x, f^*(y) \rangle$ definuje *adjungované* lineární zobrazení f^* k lineárnímu zobrazení f . V našem jednoduchém případě, kdy máme zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ a standardní skalární součin, je $\mathbf{f}^*(\mathbf{y}) = \mathbf{A}^T \mathbf{y}$, tedy adjunkce splývá s maticovou transpozicí.

⁵Výjimkou jsou např. počítačové grafici, pro které ‘vektor’ bez přívlastku znamená řádkový vektor.

- Pro $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{x} \in \mathbb{R}^m$ je výraz $\mathbf{x}^T \mathbf{A}$ maticový součin matice $1 \times m$ a matice $m \times n$, což je řádkový vektor délky n . Je to vlastně lineární kombinace řádků matice \mathbf{A} s koeficienty \mathbf{x} .
- Pro $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ je výraz

$$\mathbf{x}^T \mathbf{y} = x_1 y_1 + \cdots + x_n y_n = \langle \mathbf{x}, \mathbf{y} \rangle \quad (2.12)$$

maticový součin řádkového vektoru \mathbf{x}^T a sloupcového vektoru \mathbf{y} . Výsledek je skalár, známý jako **standardní skalární součin vektorů** \mathbf{x} a \mathbf{y} . Druhá rovnost říká (ověřte!), že je to speciální případ skalárního součinu matic (2.7), ve kterém považujeme \mathbf{x}, \mathbf{y} za matice s jedním sloupcem. Pro standardní skalární součin vektorů budeme v dalším textu vždy používat značení $\mathbf{x}^T \mathbf{y}$ a ne $\langle \mathbf{x}, \mathbf{y} \rangle$.

- Pro $\mathbf{x} \in \mathbb{R}^m$ a $\mathbf{y} \in \mathbb{R}^n$ je výraz

$$\mathbf{xy}^T = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_m y_1 & \cdots & x_m y_n \end{bmatrix} \in \mathbb{R}^{m \times n} \quad (2.13)$$

matice $m \times n$, které se někdy říká **vnější součin** vektorů \mathbf{x} a \mathbf{y} nebo **dyáda**⁶.

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 $\mathbf{1}^T \mathbf{x} = x_1 + \cdots + x_n$. Jiný příklad: jednotkovou matici můžeme značit (i když se to tak nedělá) $\mathbf{I} = \text{diag } \mathbf{1}$.

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 počet n složek vektoru \mathbf{e}_i je určen kontextem. Standardní báze tvoří sloupce jednotkové matice, $[\mathbf{e}_1 \cdots \mathbf{e}_n] = \mathbf{I}_n$.

2.3 Matice sestavené z bloků

Matici je možno sestavit z několika jejích **podmatic** (zvaných též **bloky**) jako

$$\begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \cdots & \mathbf{A}_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n} \quad (2.14)$$

kde $\mathbf{A}_{ij} \in \mathbb{R}^{m_i \times n_j}$ a $m = \sum_i m_i$ a $n = \sum_j n_j$. Např.

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad [\mathbf{A} \ \mathbf{B}], \quad \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \quad (2.15)$$

kde v prvním příkladu musí mít např. 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}, \mathbf{D} . V Matlabu matice (2.15) napíšeme jako

$$[\mathbf{A} \ \mathbf{B}; \ \mathbf{C} \ \mathbf{D}], \quad [\mathbf{A} \ \mathbf{B}], \quad [\mathbf{A}; \ \mathbf{B}], \quad [\mathbf{A} \ \text{eye}(\mathbf{m}); \ \text{zeros}(\mathbf{n}) \ \mathbf{D}] \quad (\text{kde } [\mathbf{m}, \mathbf{n}] = \text{size}(\mathbf{A})).$$

⁶Dyáda se dá považovat za tensorový součin dvou vektorů v maticové notaci. Ovšem *tensorový součin* je abstraktnější a obecnější pojem.

Při násobení matic sestavených z bloků lze neformálně užít obvyklý vzorec (2.2) pro násobení matic, ve kterém si místo prvků matice představíme bloky. Formálně,

$$\begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1p} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \cdots & \mathbf{A}_{mp} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \cdots & \mathbf{B}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{B}_{p1} & \cdots & \mathbf{B}_{pn} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \cdots & \mathbf{C}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{m1} & \cdots & \mathbf{C}_{mn} \end{bmatrix} \quad (2.16)$$

kde

$$\mathbf{C}_{ij} = \sum_{k=1}^p \mathbf{A}_{ik} \mathbf{B}_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \quad (2.17)$$

Příklady:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{AX} + \mathbf{BY} \\ \mathbf{CX} + \mathbf{DY} \end{bmatrix}, \quad [\mathbf{A} \quad \mathbf{B}] \begin{bmatrix} \mathbf{C} \\ \mathbf{D} \end{bmatrix} = \mathbf{AC} + \mathbf{BD}, \quad \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} [\mathbf{E} \quad \mathbf{F}] = \begin{bmatrix} \mathbf{AE} & \mathbf{AF} \\ \mathbf{CE} & \mathbf{CF} \end{bmatrix}.$$

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

$$\mathbf{A} = [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n], \quad (2.18)$$

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

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} = [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_m]^T, \quad (2.19)$$

kde řádkové vektory⁷ $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$ jsou řádky matice \mathbf{A} , přičemž $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$.

Vyjádríme-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

$$\mathbf{AB} = [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_p] \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_p^T \end{bmatrix} = \mathbf{a}_1 \mathbf{b}_1^T + \cdots + \mathbf{a}_p \mathbf{b}_p^T. \quad (2.20)$$

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

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} [\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n] = [\mathbf{Ab}_1 \quad \cdots \quad \mathbf{Ab}_n] = \begin{bmatrix} \mathbf{a}_1^T \mathbf{B} \\ \vdots \\ \mathbf{a}_m^T \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{b}_1 & \cdots & \mathbf{a}_1^T \mathbf{b}_n \\ \vdots & \ddots & \vdots \\ \mathbf{a}_m^T \mathbf{b}_1 & \cdots & \mathbf{a}_m^T \mathbf{b}_n \end{bmatrix}. \quad (2.21)$$

2.4 Co je soustava lineárních rovnic?

Soustava rovnic je *lineární*, jestliže se v žádné rovnici proměné nevyskytují v mocninách (např. x^2) ani v součinech (např. xy). Neboli pravá i levá strana každé rovnice je polynom více proměnných nejvýše prvního stupně. Když pojmenujeme proměnné jako x_1, \dots, x_n , každou lineární soustavu m rovnic s n neznámými jde zapsat jako

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

⁷Vektory \mathbf{a}_i samozřejmě označují jiné vektory v (2.18) a v (2.19).

neboli

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, m.$$

neboli

$$\mathbf{Ax} = \mathbf{b}, \tag{2.22}$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$. Soustava je **homogenní** pokud $\mathbf{b} = \mathbf{0}$ (tj. každé z čísel b_1, \dots, b_m je nulové), v opačném případě (aspoň jedno z čísel b_1, \dots, b_m je nenulové) je **nehomogenní**.

Chceme-li řešit lineární soustavu na počítači, příslušné algoritmy často vyžadují soustavu ve tvaru (2.22), tj. všechny neznámé jsou souřaděné do jediného vektoru. Např. v Matlabu se řešení soustavy (2.22) spočítá jednoduše jako $\mathbf{x} = \mathbf{A} \setminus \mathbf{b}$ (zde předpokládáme, že soustava má právě jedno řešení). Ovšem ne vždy dostaneme lineární soustavu v tomto tvaru. Přepisujeme-li takové soustavy do tvaru (2.22), můžeme snadno udělat chybu. Existují způsoby, jak to dělat elegantněji a obecněji⁸, viz následující příklady a Cvičení 2.6.

Příklad 2.1. Soustava

$$\mathbf{Ax} = \mathbf{b}, \tag{2.23a}$$

$$\mathbf{A}^T \mathbf{y} = \mathbf{x}, \tag{2.23b}$$

kde $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{y} \in \mathbb{R}^m$ jsou neznámé a $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$ jsou dány. Je to lineární soustava $m+n$ rovnic s $m+n$ neznámými. Tuto soustavu můžeme napsat ve tvaru (2.22) (kde samozřejmě písmena \mathbf{A} , \mathbf{b} znamenají něco jiného než v naší soustavě) takto:

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{I} & -\mathbf{A}^T \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}.$$

Samořejmě, že soustavu (2.23) lze vyřešit i jinak. Dosadíme \mathbf{x} z druhé rovnice do první, což dá $\mathbf{AA}^T \mathbf{y} = \mathbf{b}$. Tuto rovnici vyřešíme pro \mathbf{y} (pokud má řešení) a spočítáme \mathbf{x} z druhé rovnice. ♦

Příklad 2.2. Homogenní lineární soustava

$$\mathbf{a}_i^T \mathbf{x} = \alpha_i b_i \quad \forall i = 1, \dots, m$$

kde $\mathbf{x} \in \mathbb{R}^n$ a $\alpha_1, \dots, \alpha_m$ jsou neznámé a $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a $b_1, \dots, b_m \in \mathbb{R}$ jsou dány. Označíme-li $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_m]^T \in \mathbb{R}^{m \times n}$, $\mathbf{b} = (b_1, \dots, b_m)$ a $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$, pak soustavu můžeme psát jako $\mathbf{Ax} = (\text{diag } \mathbf{b})\boldsymbol{\alpha}$, neboli

$$[\mathbf{A} \quad -\text{diag } \mathbf{b}] \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\alpha} \end{bmatrix} = \mathbf{0}. \quad \blacklozenge$$

⁸Abychom řekli přesně, co znamená ‘obecněji’, musíme rozlišovat mezi *instancí* lineární soustavy, což je jedna konkrétní soustava, a *nekonečnou množinou instancí* lineárních soustav v určitém tvaru (kterou někdy krátce nazýváme *problém* nebo *úloha*; to jsou pojmy z *teorie složitosti algoritmů*). Např. soustava $\{x - 2y = 1, -x + y = 3\}$ je instance, ale (2.23) je nekonečná množina lineárních soustav, jejíž každý prvek je určen konkrétní dvojicí (\mathbf{A}, \mathbf{b}) . Převodem úlohy (2.23) do tvaru (2.22) pak rozumíme *algoritmus*, který pro libovolnou dvojici (\mathbf{A}, \mathbf{b}) najde dvojici $(\mathbf{A}', \mathbf{b}')$ tak, že soustava (2.23) bude ekvivalentní soustavě $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$ kde $\mathbf{x}' = (\mathbf{x}, \mathbf{y})$.

Příklad 2.3. Soustava

$$\sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m, \quad (2.24a)$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n, \quad (2.24b)$$

kde $x_{11}, \dots, x_{mn} \in \mathbb{R}$ jsou neznámé a $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{R}$ jsou dány. Je to lineární soustava $m + n$ rovnic s mn neznámými. Tato soustava jde napsat v maticovém tvaru jako

$$\begin{aligned} \mathbf{X}\mathbf{1} &= \mathbf{a}, \\ \mathbf{X}^T\mathbf{1} &= \mathbf{b}, \end{aligned}$$

kde $\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{R}^m$, $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$ a $\mathbf{X} = [x_{ij}] \in \mathbb{R}^{m \times n}$. To ale není tvar (2.22), neboť proměnné jsou opět soustředěné do matice \mathbf{X} a nikoliv do vektoru. Soustavu ale lze přepsat takto (ověřte roznásobením matic!):

$$\begin{bmatrix} \mathbf{I}_m & \cdots & \mathbf{I}_m \\ \mathbf{1}_m^T & & \\ & \ddots & \\ & & \mathbf{1}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \quad (2.25)$$

kde $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^m$ jsou sloupce matice \mathbf{X} . ◆

2.5 Maticové zločiny

Při manipulaci s maticovými výrazy a rovnicemi dělají někteří studenti hrubé chyby, kterých se lze při alespoň minimální snaze vyhnout. Takové chyby jsou *neomluvitelné*. Uvedme typické příklady těchto zločinů.

Výraz je nesmyslný kvůli rozměrům matic

Jako první 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 $\mathbf{B}^{-1}\mathbf{A}$.

Abyste měli jasnou představu o rozměrech matic v maticových výrazech, pod matice a vektory si malujte obdélníčky s rozměry matic!

- Předpoklad, že existuje pravá inverze úzké matice, např. $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$ pro $\mathbf{Q} \in \mathbb{R}^{5 \times 3}$.
- Inverze čtvercové, ale evidentně singulární matice, např. $(\mathbf{A}\mathbf{B})^{-1}$ kde \mathbf{A} je úzká. Kdyby totiž inverze existovala, bylo by $\mathbf{A}\mathbf{B}\mathbf{B}^T\mathbf{A} = \mathbf{I}$, tedy \mathbf{A} by měla pravou inverzi. Ale to nelze, protože \mathbf{A} je úzká. Extrémnější příklad stejného zločinu je $(\mathbf{w}\mathbf{w}^T)^{-1}$ kde \mathbf{w} je vektor.

Příklad 2.4. Vidíme-li výraz $(\mathbf{A}^T \mathbf{B})^{-1}$, musí nám ve zlomku vteřiny hlavou proběhnout tyto úvahy o rozměrech matic:

- Aby matice bylo možné násobit, musí mít stejný počet řádků.
- Protože nelze invertovat obdélníkovou matici, musí být součin $\mathbf{A}^T \mathbf{B}$ čtvercový. Tedy obě matice musí mít stejný počet sloupců. Ted' víme, že obě matice musí mít stejný rozměr.
- Pokud by \mathbf{A}^T byla úzká nebo \mathbf{B} široká, $\mathbf{A}^T \mathbf{B}$ by byla singulární a tedy by neměla inverzi. Tedy musí být obě matice čtvercové nebo úzké. \blacklozen

Abyste nedělali chyby v rozměrech matic, pod maticové výrazy si pište rozměry matic, např.

$$\underbrace{\underbrace{\mathbf{A}^T}_{m \times n} \underbrace{\mathbf{B}}_{m \times n}}_{n \times n}.$$

Použití neexistujících maticových identit

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

- $(\mathbf{AB})^T = \mathbf{A}^T \mathbf{B}^T$ (pokud v maticovém součinu $\mathbf{A}^T \mathbf{B}^T$ je vnitřní rozměr různý, je to chyba už kvůli rozměrům matic)
- $(\mathbf{AB})^{-1} = \mathbf{A}^{-1} \mathbf{B}^{-1}$ (podobná situace)
- $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$. Tato identita vyplývá z neexistující (avšak velice užitečné) identity $\mathbf{AB} = \mathbf{BA}$. Správně je $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2$.

Pracujte nejen s papírem, ale i s Matlabem! Hypotézy o platnosti či neplatnosti maticových rovnic lze často testovat na náhodných maticích. Např. nejste-li si jisti rovností $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$, zkuste `A=randn(5,3)`; `B=randn(3,6)`; `(A*B)' - B'*A'` (což ale samozřejmě není důkaz).

Neekvivalentní úpravy rovnic

Zde pachatel udělá *neekvivalentní úpravu* rovnice, ale myslí si, že udělal ekvivalentní. 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 $a^2 = b^2 \not\Rightarrow a = b$. 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 nenulový).
- 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).

Soustavy rovnic (lineárních či nelineárních) se často snažíme řešit (nebo aspoň zjednodušit) eliminací některých proměnných. To obvykle znamená, že v některé rovnici osamostatníme jednu proměnnou a dosadíme za tuto proměnnou do ostatních rovnic. Např. ze soustavy

$$x^2 + y^2 = 1, \quad x - 2y = 1 \tag{2.26}$$

eliminujeme proměnnou x tak, že druhou rovnici přepíšeme na $x = 2y + 1$ a pravou stranu dosadíme do první, což dá

$$(2y + 1)^2 + y^2 = 1. \tag{2.27}$$

Ovšem je nutné, aby upravená soustava byla ekvivalentní původní soustavě, tj. pro všechna $x, y \in \mathbb{R}$ musí platit (2.26) \Leftrightarrow (2.27). To je pravda, protože $x - 2y = 1 \Leftrightarrow x = 2y + 1$. Pokud bychom ale chtěli “vyjádřit” x z první rovnice jako $x = \sqrt{1 - y^2}$ a dosadit do druhé rovnice, sice $x = \sqrt{1 - y^2} \Rightarrow x^2 + y^2 = 1$ ale $x^2 + y^2 = 1 \not\Rightarrow x = \sqrt{1 - y^2}$. Následkem toho by výsledná rovnice $\sqrt{1 - y^2} - 2y = 1$ nebyla ekvivalentní původní soustavě (2.26).

Tohle dobře znáte ze skalárního algebry, ale ne tak dobře v maticové algebře. Student řeší např. soustavu rovnic

$$\mathbf{a}^T \mathbf{x} = 0, \quad \mathbf{x}^T \mathbf{x} = 1$$

(kde $\mathbf{a} \in \mathbb{R}^n$ je dáno a $\mathbf{x} \in \mathbb{R}^n$ je neznámá) tak, že se snaží “vyjádřit” \mathbf{x} z první rovnice a dosadit ho do druhé. To je těžký zločin, protože rovnice $\mathbf{a}^T \mathbf{x} = 0$ má nekonečně mnoho řešení (pro $n > 1$) a tedy z ní neplyne, že \mathbf{x} je rovno jakémukoliv jednomu vektoru.

2.6 Cvičení

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

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

2.2. Řešíme soustavu rovnic $\mathbf{b}_i = \mathbf{Xa}_i$ (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. Chceme vyřešit soustavu rovnic

$$\begin{aligned} \mathbf{Ax} + (\mathbf{y}^T \mathbf{B})^T &= \alpha \mathbf{1} \\ \mathbf{Ay} + \mathbf{c} &= \mathbf{0} \end{aligned}$$

kde \mathbf{A} , \mathbf{B} jsou známé matice, \mathbf{c} je známý vektor, \mathbf{x} , \mathbf{y} jsou neznámé vektory a α je neznámý skalár. Soustavu přepište do tvaru $\mathbf{Pu} = \mathbf{q}$, kde matice \mathbf{P} a vektor \mathbf{q} obsahují známé konstanty a vektor \mathbf{u} obsahuje všechny neznámé.

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

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

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

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

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

2.6. Zobrazení $\text{vec}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{mn}$ (vektORIZACE matice, v Matlabu označeno $\mathbf{A}(:)$) je definováno tak, že vektor $\text{vec } \mathbf{A} \in \mathbb{R}^{mn}$ je matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ přerovnaná po sloupcích do vektoru. Např. $\text{vec} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = (1, 3, 2, 4) = [1 \ 3 \ 2 \ 4]^T$. *Kroneckerův součin* dvou matic je definován jako

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

v Matlabu $\text{kron}(\mathbf{A}, \mathbf{B})$. Pro libovolné matice (s kompatibilními velikostmi) platí

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec } \mathbf{B}. \quad (2.28)$$

Použijte tohoto vzorce pro transformaci následujících soustav rovnic s neznámou maticí \mathbf{X} do tvaru (2.22). Předpokládejte, že matice a vektory mají nejobecnější možné rozměry, které jsou kompatibilní s maticovými operacemi v soustavě.

- a) Soustava $\mathbf{b}_i^T \mathbf{Xa}_i = 0 \ \forall i = 1, \dots, k$, kde neznámá je matice \mathbf{X} .
- b) Soustava $\mathbf{AX} + \mathbf{X}^T \mathbf{B}^T = \mathbf{C}$, kde neznámá je matice \mathbf{X} .
- c) Soustava z Příkladu 2.3.

2.7. *Komutátor* dvou matic je matice $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. 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) $[\mathbf{A}, \mathbf{BC}] = [\mathbf{A}, \mathbf{B}]\mathbf{C} + \mathbf{B}[\mathbf{A}, \mathbf{C}]$

2.8. Pro jaké matice platí, že jsou samy sobě inverzí? Odpověď dokažte.

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

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

2.11. Za předpokladu, že invertované matice jsou invertovatelné, dokažte identity:

- a) $\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B} = (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1} = \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{A}$
- b) $(\mathbf{I} + \mathbf{AB})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I} + \mathbf{BA})^{-1}\mathbf{B}$
- c) $(\mathbf{I} + \mathbf{AB})^{-1}\mathbf{A} = \mathbf{A}(\mathbf{I} + \mathbf{AB})^{-1}$
- d) $\mathbf{A} - \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{A} = \mathbf{B} - \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}$

2.12. Dokažte

- a) *vzorec Shermana a Morrisonové*

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

za předpokladu, že \mathbf{A} je regulární a $\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} + 1 \neq 0$.

b) vzorec Shermana, Morrisonové a Woodburyho neboli lema o maticové inverzi

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

za předpokladu, že \mathbf{A} , \mathbf{C} a $\mathbf{C}^{-1} + \mathbf{V}^T\mathbf{A}^{-1}\mathbf{U}$ jsou regulární.

c) Vzorec pro inverzi blokové matice

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{Q}^{-1} & -\mathbf{Q}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{Q}^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{Q}^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix} \quad \text{kde } \mathbf{Q} = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$$

za předpokladu, že \mathbf{D} a \mathbf{Q} jsou regulární. (Matice \mathbf{Q} je známa jako *Schurův doplněk* invertované blokové matice.) Srov. Cvičení 2.4.

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

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

2.15. Jistý student tvrdí, že když matice \mathbf{A} má levou inverzi (označme ji \mathbf{B}), tak lineární soustavu $\mathbf{Ax} = \mathbf{b}$ lze vyřešit jednoduše vynásobením zleva maticí \mathbf{B} . Dostaneme $\mathbf{BAx} = \mathbf{Bb}$, a jelikož $\mathbf{BA} = \mathbf{I}$, máme tedy $\mathbf{x} = \mathbf{Bb}$. Má student pravdu? Vysvětlete.

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

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

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

2.18. 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*.

2.19. Dokažte, že rovnice $\mathbf{XY} - \mathbf{YX} = \mathbf{I}$ nemá řešení pro žádné \mathbf{X}, \mathbf{Y} .

2.20. Dokažte (2.10) bez použití (2.8), tj. přímo z definic (2.7) a (2.2).

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.
$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^T & \mathbf{1} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ -\alpha \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\mathbf{c} \end{bmatrix}$$

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.

2.6.a) Je $\mathbf{b}_i^T \mathbf{X} \mathbf{a}_i = \text{vec}(\mathbf{b}_i^T \mathbf{X} \mathbf{a}_i) = (\mathbf{a}_i^T \otimes \mathbf{b}_i^T) \text{vec } \mathbf{X}$. Tedy soustavu lze napsat jako $\begin{bmatrix} \mathbf{a}_1^T \otimes \mathbf{b}_1^T \\ \vdots \\ \mathbf{a}_k^T \otimes \mathbf{b}_k^T \end{bmatrix} \text{vec } \mathbf{X} = \mathbf{0}$.

2.6.b) Musí být $\mathbf{X} \in \mathbb{R}^{m \times n}$ je neznámá a $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times m}$ a $\mathbf{C} \in \mathbb{R}^{n \times n}$ jsou dány. Jedná se tedy o lineární soustavu n^2 rovnic s mn neznámými. Neznámé ovšem nejsou soustředěné do vektoru, ale do matice, tedy soustava není ve tvaru (2.22). Je $\text{vec}(\mathbf{A}\mathbf{X}) = \text{vec}(\mathbf{A}\mathbf{X}\mathbf{I}_n) = (\mathbf{I}_n \otimes \mathbf{A}) \text{vec } \mathbf{X}$ a $\text{vec}(\mathbf{X}\mathbf{B}^T) = \text{vec}(\mathbf{I}_m \mathbf{X}\mathbf{B}^T) = (\mathbf{B} \otimes \mathbf{I}_m) \text{vec } \mathbf{X}$. Tedy soustavu můžeme napsat jako $((\mathbf{I}_n \otimes \mathbf{A}) + (\mathbf{B} \otimes \mathbf{I}_m)) \text{vec } \mathbf{X} = \text{vec } \mathbf{C}$.

2.6.c) Je $\text{vec}(\mathbf{X}\mathbf{1}_n) = \text{vec}(\mathbf{I}_m \mathbf{X}\mathbf{1}_n) = (\mathbf{1}_n^T \otimes \mathbf{I}_m) \text{vec } \mathbf{X}$ a $\text{vec}(\mathbf{1}_m^T \mathbf{X}) = \text{vec}(\mathbf{1}_m \mathbf{X}\mathbf{I}_n) = (\mathbf{I}_n \otimes \mathbf{1}_m) \text{vec } \mathbf{X}$. Tedy soustavu můžeme psát jako $\begin{bmatrix} \mathbf{1}_n^T \otimes \mathbf{I}_m \\ \mathbf{I}_n \otimes \mathbf{1}_m \end{bmatrix} \text{vec } \mathbf{X} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$. Zkontrolujte vypočítáním výrazů $\mathbf{1}_n^T \otimes \mathbf{I}_m$ a $\mathbf{I}_n \otimes \mathbf{1}_m$, že je to totéž jako soustava (2.25).

2.7.a) $[\mathbf{A}, \mathbf{B}]^T = (\mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A})^T = \mathbf{B}^T \mathbf{A}^T - \mathbf{A}^T \mathbf{B}^T = \mathbf{B}\mathbf{A} - \mathbf{A}\mathbf{B} = -[\mathbf{A}, \mathbf{B}]$

2.7.c) Máme dokázat, že $\mathbf{A}\mathbf{B}\mathbf{C} - \mathbf{B}\mathbf{C}\mathbf{A} = (\mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A})\mathbf{C} + \mathbf{B}(\mathbf{A}\mathbf{C} - \mathbf{C}\mathbf{A})$, což zjevně platí.

2.9. Z definice inverze stačí ukázat, že $\mathbf{A}\mathbf{B}\mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{I}$. To je očividné, protože $\mathbf{B}\mathbf{B}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

2.11.a) Stačí dokázat první rovnost, druhá platí symetricky. Máme ukázat $\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}(\mathbf{A}^{-1} + \mathbf{B}^{-1}) = \mathbf{I}$. Je $\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}(\mathbf{A}^{-1} + \mathbf{B}^{-1}) = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}\mathbf{A}^{-1} + \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{B}\mathbf{A}^{-1} + \mathbf{I})$. Vynásobením rovnosti $\mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{B}\mathbf{A}^{-1} + \mathbf{I}) = \mathbf{I}$ zleva maticí $(\mathbf{A} + \mathbf{B})\mathbf{A}^{-1}$ (což je ekvivalentní úprava) dá $(\mathbf{B}\mathbf{A}^{-1} + \mathbf{I}) = (\mathbf{A} + \mathbf{B})\mathbf{A}^{-1}$, což zjevně platí.

2.11. Necht' \mathbf{A} je regulární a symetrická (tedy čtvercová). Necht' $\mathbf{B}\mathbf{A} = \mathbf{I}$, tedy \mathbf{B} je levá inverze \mathbf{A} . Z toho $\mathbf{A}^T \mathbf{B}^T = \mathbf{A}\mathbf{B}^T = \mathbf{I}$, tedy \mathbf{B}^T je pravá inverze \mathbf{A} . Ale pro regulární matici jsou si levá a pravá inverze rovny (viz §2.1.4), tedy $\mathbf{B}^T = \mathbf{B}$.

2.12.a) Musíme ukázat, že $(\mathbf{A} + \mathbf{u}\mathbf{v}^T) \left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \right) = \mathbf{I}$. To je pouhé cvičení na úpravy maticových výrazů, ale nesmíme se do toho zamotat a udělat chybu. Upravujeme:

$$\begin{aligned} (\mathbf{A} + \mathbf{u}\mathbf{v}^T) \left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \right) &= \mathbf{I} - \frac{\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \frac{\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \\ &= \mathbf{I} - \frac{\mathbf{u}\mathbf{w}^T}{1 + \mathbf{w}^T\mathbf{u}} + \mathbf{u}\mathbf{w}^T - \frac{\mathbf{u}\mathbf{w}^T\mathbf{u}\mathbf{w}^T}{1 + \mathbf{w}^T\mathbf{u}} \\ &= \mathbf{I} - \frac{\mathbf{u}\mathbf{w}^T - \mathbf{u}\mathbf{w}^T - (\mathbf{w}^T\mathbf{u})\mathbf{u}\mathbf{w}^T + \mathbf{u}\mathbf{w}^T\mathbf{u}\mathbf{w}^T}{1 + \mathbf{w}^T\mathbf{u}}, \end{aligned}$$

kde jsme označili $\mathbf{v}^T\mathbf{A}^{-1} = \mathbf{w}^T$. Výraz $\mathbf{u}\mathbf{w}^T\mathbf{u}\mathbf{w}^T$ je maticový součin čtyř matic. Ale výraz $\mathbf{w}^T\mathbf{u}$ je skalár, proto $\mathbf{u}\mathbf{w}^T\mathbf{u}\mathbf{w}^T = (\mathbf{w}^T\mathbf{u})\mathbf{u}\mathbf{w}^T$. Tedy čítec zlomku je nulový vektor a jsme hotovi.

2.13. Matice $\text{diag}(a_1, \dots, a_n)$ je regulární, právě když a_1, \dots, a_n jsou všechny nenulové. Je $\text{diag}(a_1, \dots, a_n)^{-1} = \text{diag}(1/a_1, \dots, 1/a_n)$. Obojí plyne přímo z definic regularity, inverze a diagonální matice.

2.14. Plyne snadno z definice diagonální matice a maticového součinu.

- 2.15. Nápověda: Nemá pravdu, protože (za podmínky $\mathbf{BA} = \mathbf{I}$) sice platí $\mathbf{Ax} = \mathbf{b} \Rightarrow \mathbf{x} = \mathbf{Bb}$, ale neplatí $\mathbf{Ax} = \mathbf{b} \Leftarrow \mathbf{x} = \mathbf{Bb}$. Zkuste si to třeba pro $\mathbf{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ a $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.
- 2.16.a) Máme dokázat, že $(\mathbf{A} + \mathbf{A}^T)^T = \mathbf{A} + \mathbf{A}^T$. To je jasné, protože $(\mathbf{A}^T)^T = \mathbf{A}$ a tedy $(\mathbf{A} + \mathbf{A}^T)^T = \mathbf{A}^T + \mathbf{A}$.
- 2.16.b) Analogické jako minule.
- 2.16.c) Takové matice existují, konkrétně $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$ a $\mathbf{C} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T)$. Nahoře jsme ukázali, že $\mathbf{B}^T = \mathbf{B}$ a $\mathbf{C}^T = -\mathbf{C}$. Jednoznačnost dokážeme takto: předpokládejme, že pro symetrické matice \mathbf{B}, \mathbf{B}' a antisymetrické \mathbf{C}, \mathbf{C}' platí $\mathbf{B} + \mathbf{C} = \mathbf{B}' + \mathbf{C}'$. Z toho plyne $\mathbf{B}^T + \mathbf{C}^T = \mathbf{B}'^T + \mathbf{C}'^T$, tedy $\mathbf{B} - \mathbf{C} = \mathbf{B}' - \mathbf{C}'$. Porovnáním s předchozí rovnicí dostaneme $\mathbf{B} = \mathbf{B}'$ a $\mathbf{C} = \mathbf{C}'$.
- 2.16.d) Plyne okamžitě z identity $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$ (viz §2.1.3).
- 2.18. Opět pouhé cvičení na úpravy maticových výrazů.
- 2.19. Použijte cykličnost stopy.

Kapitola 3

Linearita

Množina \mathbb{R}^n spolu s operacemi sčítání vektorů a násobení vektorů skalárem tvoří *lineární prostor* (také zvaný *vektorový prostor*)¹ nad tělesem \mathbb{R} . Protože budeme pracovat pouze v prostoru \mathbb{R}^n (a občas v prostoru $\mathbb{R}^{m \times n}$), nebudeme uvádět definici (pomocí axiomů) lineárního prostoru ani odkazovat se na ni – určitě si ji ale najdete a zopakujte! To nám sice občas zabránilo využít celou sílu a krásu lineární algebry, ale zkrátí a zjednoduší to výklad.

3.1 Podprostory

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 $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou **lineárně nezávislé**², jestliže

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

V opačném případě (tedy když pro nějaká $\alpha_1, \dots, \alpha_k$, aspoň jedno z nich nenulové, platí $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \mathbf{0}$) jsou vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ **lineárně závislé**. Lze dokázat (viz Cvičení 3.4) že jestliže jsou vektory lineárně závislé, tak je aspoň jeden z nich lineární kombinací ostatních.

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

Důkaz. Necht' kromě (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 $\alpha_i = \beta_i$. ■

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

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

¹Jaký je vlastně rozdíl mezi *množinou* a *prostorem*? Množina je ‘pytel’ prvků (‘věcí’) bez jakékoliv struktury. Prostor je neformální název pro *algebraickou strukturu*, která má nějaký geometrický význam. Algebraická struktura je množina (někdy i více množin) spolu s *operacemi* definovanými na této množině (příklady: grupa, těleso, lineární prostor, lineární prostor se skalárním součinem). Tyto operace a jejich vlastnosti určují vlastnosti algebraické struktury a tedy i prostoru. Operace se ale v matematickém textu často explicitně nezmiňují (je z kontextu jasné, které to jsou) a tak se místo o algebraické strukturu mluví jen o množině.

²Tato formulace je trochu nepřesná, protože lineární nezávislost není vlastnost jednotlivých vektorů, ale posloupnosti (či seznamu) vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$. Přesto je široce používána.

všech jejich lineárních kombinací. Zde mlčky předpokládáme, že vektorů je konečný počet (lineární obal nekonečné množiny vektorů se musí definovat jinak, více o tom později v §13.2).

Neprázdna množina $X \subseteq \mathbb{R}^n$ se nazývá **lineární podprostor**³ (nebo jen **podprostor**) lineárního prostoru \mathbb{R}^n , jestliže každá lineární kombinace každé (konečné) množiny vektorů z X leží v X (neboli množina X je uzavřená na lineární kombinace):

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

Tvrzení (3.3) je ekvivalentní dvěma tvrzeními

$$\mathbf{x}, \mathbf{y} \in X \implies \mathbf{x} + \mathbf{y} \in X, \quad (3.4a)$$

$$\mathbf{x} \in X, \alpha \in \mathbb{R} \implies \alpha \mathbf{x} \in X, \quad (3.4b)$$

kde implikace (3.3) \implies (3.4) je očividná a implikace (3.4) \implies (3.3) by se dokázala např. indukcí. Snadno se ukáže, že lineární obal libovolné množiny vektorů je lineární podprostor.

Každý podprostor je neprázdny, protože sám o sobě to musí být lineární *prostor*, a ten ze své definice musí obsahovat nulový vektor $\mathbf{0}$ (kterému se z historických důvodů také říká *počátek*). Podprostor $X = \{\mathbf{0}\}$ se nazývá **triviální**.

Báze lineárního podprostoru $X \subseteq \mathbb{R}^n$ je lineárně nezávislá množina⁴ vektorů, jejichž lineární obal je X . Platí tato zásadní tvrzení (důkazy nejsou krátké a neuvádíme je, najdete je v každé učebnici lineární algebry):

Věta 3.2 (vlastnosti báze).

- Každý lineární podprostor má (alespoň jednu) bázi.
- Každá báze lineárního podprostoru má stejný počet vektorů.
- Z každé množiny vektorů lze vybrat bázi jejich lineárního obalu.
- Každou lineárně nezávislou množinu vektorů z lineárního podprostoru lze doplnit na jeho bázi.

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 $\mathbf{x} \in X$, pak (jednoznačně určené) skaláry $\alpha_1, \dots, \alpha_k$ v (3.1) se nazývají **souřadnice** vektoru \mathbf{x} v bázi $\mathbf{x}_1, \dots, \mathbf{x}_k$.

Tvrzení 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. Jestliže $X \subseteq Y$, 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 bázi Y (tedy doplnění nepřidá žádný vektor), odtud druhé tvrzení. ■

Příklad 3.1. Uved' me jednoduché příklady podprostorů prostorů \mathbb{R}^n :

³Krátce budeme říkat, že $X \subseteq \mathbb{R}^n$ je podprostor. To nemůže způsobit zmatení, neboť aby množina byla podprostor prostoru \mathbb{R}^n , musí všechny její prvky patřit do \mathbb{R}^n . Např. množina $X \subseteq \mathbb{R}^3$ tedy nemůže být podprostor prostoru \mathbb{R}^2 , to by byl nesmysl.

⁴Někdy (např. když mluvíme o souřadnicích vektoru vzhledem k bázi) nám záleží na pořadí prvků báze. Pak musíme bázi považovat ne za množinu, ale za *uspořádanou* množinu nebo *posloupnost* vektorů.

- Triviálně, množina $X = \mathbb{R}^3$ je podprostorem prostoru \mathbb{R}^3 . Jeho dimenze je 3. Jeho báze je např. množina $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ (standardní báze) nebo $\{(1, 1, 1), (1, -1, 0), (2, 0, 0)\}$.
- Množina $X = \text{span}\{(1, 2, 3)\} = \{\alpha(1, 2, 3) \mid \alpha \in \mathbb{R}\} \subseteq \mathbb{R}^3$ je podprostor dimenze 1 prostoru \mathbb{R}^3 . Jeho báze je např. množina $\{(1, 2, 3)\}$, jiná báze je $\{(2, 4, 6)\}$. Je to přímka procházející počátkem.
- Množina $X = \text{span}\{(1, 2, 3), (1, 0, -1)\} = \{\alpha(1, 2, 3) + \beta(1, 0, -1) \mid \alpha, \beta \in \mathbb{R}\}$ je podprostor dimenze 2 prostoru \mathbb{R}^3 . Jeho báze je např. množina $\{(1, 2, 3), (1, 0, -1)\}$, jiná báze je $\{(3, 2, 1), (0, 2, 4)\}$. Je to rovina procházející počátkem. Všimněte si, že tato množina nemá žádnou ‘přirozenou’ nebo ‘standardní’ bázi jako mají prostory \mathbb{R}^n .
- Množina všech trojic $(x_1, x_2, x_3) \in \mathbb{R}^3$ splňujících rovnici $x_1 + 2x_2 - x_3 = 0$ je podprostor dimenze 2 prostoru \mathbb{R}^3 . Jeho báze je libovolná dvojice lineárně nezávislých vektorů $\{(a_1, a_2, a_3), (b_1, b_2, b_3)\}$ splňujících $a_1 + 2a_2 - a_3 = 0$ a $b_1 + 2b_2 - b_3 = 0$, např. $\{(1, 0, 1), (0, 1, 2)\}$. Je to rovina procházející počátkem.
- 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 $\{(1, 2, 3, 0), (0, 1, 2, -1)\}$.
- Všechny možné podprostory prostoru \mathbb{R}^3 jsou 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).
- Množina $X = \{(1+\alpha, \alpha) \mid \alpha \in \mathbb{R}\} \subseteq \mathbb{R}^2$ (přímka neprocházející počátkem) není podprostor, protože např. $(1, 0) \in X$ ale $2(1, 0) = (2, 0) \notin X$. \blacklozen

3.1.1 Komplementární podprostory a projekce

Každý podprostor $X \subseteq \mathbb{R}^n$ obsahuje nulový vektor, tedy $\mathbf{0} \in X$. Zajímavá je situace, kdy dva podprostory $X, X' \subseteq \mathbb{R}^n$ mají minimální možný průnik, tedy⁵ $X \cap X' = \{\mathbf{0}\}$. Takové dvojice podprostorů přirozeně vznikají tak, že nějakou bázi rozložíme na dvě podmnožiny:

Tvrzení 3.4. Necht' $\mathbf{x}_1, \dots, \mathbf{x}_m$ je báze podprostoru X a $\mathbf{x}_{m+1}, \dots, \mathbf{x}_k$ je báze podprostoru X' . Pak $X \cap X' = \{\mathbf{0}\}$ právě tehdy, když vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou lineárně nezávislé.

Důkaz. Je $\mathbf{x} \in X \cap X'$ právě tehdy, když pro nějaká $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ platí

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_m \mathbf{x}_m = -\alpha_{m+1} \mathbf{x}_{m+1} - \dots - \alpha_k \mathbf{x}_k. \quad (3.5)$$

Jestliže $X \cap X' = \{\mathbf{0}\}$, pak nutně $\mathbf{x} = \mathbf{0}$ a tedy (protože $\mathbf{x}_1, \dots, \mathbf{x}_m$ a $\mathbf{x}_{m+1}, \dots, \mathbf{x}_k$ jsou báze) $\alpha_1 = \dots = \alpha_k = 0$. Dle (3.2) jsou tedy $\mathbf{x}_1, \dots, \mathbf{x}_k$ lineárně nezávislé.

Obráceně, jestliže $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou lineárně nezávislé, pak podle (3.2) rovnost (3.5) implikuje $\alpha_1 = \dots = \alpha_k = 0$ a tedy $X \cap X' = \{\mathbf{0}\}$. \blacksquare

Ještě zajímavější situace nastane, když podprostory X, X' mají minimální průnik a zároveň jsou ‘největší možné’. Dvojici podprostorů $X, X' \subseteq \mathbb{R}^n$ nazveme **komplementární**, jestliže⁶

$$X \cap X' = \{\mathbf{0}\}, \quad (3.6a)$$

$$\dim X + \dim X' = n. \quad (3.6b)$$

⁵Jistě chápete, že $\{\mathbf{0}\}$ neoznačuje prázdnou množinu, ale množinu obsahující pouze nulový vektor.

⁶Kdo zná více z lineární algebry, rozpozná zde *direktní součet* podprostorů $X \oplus X' = \mathbb{R}^n$.

Příklad 3.2. V \mathbb{R}^2 jsou každé dvě přímky, které procházejí počátkem a protínají se pouze v počátku, komplementární podprostory. V \mathbb{R}^3 jsou přímka a rovina, které se protínají pouze v počátku, komplementární podprostory. V \mathbb{R}^3 nejsou žádné dvě přímky ani žádné dvě roviny komplementární. Rovina v \mathbb{R}^3 a přímka ležící v této rovině nejsou komplementární. \blacklozen

Věta 3.5. Necht' $X, X' \subseteq \mathbb{R}^n$ jsou komplementární podprostory. Pak pro každý vektor $\mathbf{y} \in \mathbb{R}^n$ existuje právě jedna dvojice vektorů $\mathbf{x} \in X$ a $\mathbf{x}' \in X'$ taková, že $\mathbf{y} = \mathbf{x} + \mathbf{x}'$.

Důkaz. Označme $\dim X = m$ a $\dim X' = n - m$. Necht' $\mathbf{x}_1, \dots, \mathbf{x}_m$ je báze podprostoru X a $\mathbf{x}_{m+1}, \dots, \mathbf{x}_n$ je báze podprostoru X' . Dle Tvzení 3.4 jsou vektory $\mathbf{x}_1, \dots, \mathbf{x}_n$ lineárně nezávislé, tedy tvoří bázi \mathbb{R}^n . Proto každý vektor $\mathbf{y} \in \mathbb{R}^n$ lze napsat jako

$$\mathbf{y} = \underbrace{\alpha_1 \mathbf{x}_1 + \dots + \alpha_m \mathbf{x}_m}_{\mathbf{x} \in X} + \underbrace{\alpha_{m+1} \mathbf{x}_{m+1} + \dots + \alpha_n \mathbf{x}_n}_{\mathbf{x}' \in X'}. \quad (3.7)$$

Dokažme jednoznačnost rozkladu. Necht' nějaká $\mathbf{x}_1, \mathbf{x}_1 \in X$ a $\mathbf{x}'_2, \mathbf{x}'_2 \in X'$ splňují $\mathbf{x}_1 + \mathbf{x}'_1 = \mathbf{x}_2 + \mathbf{x}'_2$, tedy $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{x}'_1 - \mathbf{x}'_2$. Protože X a X' jsou podprostory, je $\mathbf{x}_1 - \mathbf{x}_2 \in X$ a $\mathbf{x}'_1 - \mathbf{x}'_2 \in X'$, tedy $\mathbf{x}_1 - \mathbf{x}_2 \in X \cap X'$. Protože $X \cap X' = \{\mathbf{0}\}$, je $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, tedy $\mathbf{x}_1 = \mathbf{x}_2$ a $\mathbf{x}'_1 = \mathbf{x}'_2$. \blacksquare

Vektor \mathbf{x} ve Větě 3.5 se nazývá **projekce** vektoru \mathbf{y} na podprostor X ve směru podprostoru X' . Symetricky, \mathbf{x}' je projekce \mathbf{y} na X' ve směru X (viz obrázek v §3.2.5). Vzorečky pro \mathbf{x} a \mathbf{x}' bychom získali tak, že bychom spočítali souřadnice $\alpha_1, \dots, \alpha_n$ z (3.7) a dosadili zpátky. To ale dělat nebudeme, tyto vzorečky lépe vyplynou až ze §3.2.5, kde si povíme o maticích reprezentujících projekce neboli *projektorech*. Speciální případ projekce je *ortogonální projekce*, popsána v §4.6. Projekce, která není ortogonální, se nazývá *šikmá*.

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

$$\mathbf{f}(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) = \alpha_1 \mathbf{f}(\mathbf{x}_1) + \dots + \alpha_k \mathbf{f}(\mathbf{x}_k). \quad (3.8)$$

Tato podmínka je ekvivalentní dvěma podmínkám, pro každé $\mathbf{x}, \mathbf{y} \in X$ a $\alpha \in \mathbb{R}$,

$$\mathbf{f}(\mathbf{x} + \mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{y}), \quad (3.9a)$$

$$\mathbf{f}(\alpha \mathbf{x}) = \alpha \mathbf{f}(\mathbf{x}). \quad (3.9b)$$

Věta 3.6. Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je lineární, právě když existuje matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ splňující⁷

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} \quad (3.10)$$

pro všechna $\mathbf{x} \in \mathbb{R}^n$. V tom případě je matice \mathbf{A} zobrazením \mathbf{f} určena jednoznačně.

⁷Možná se ptáte, proč tedy nedefinujeme lineární zobrazení rovnou vztahem (3.10). Pro zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je to opravdu možné a definici (3.8) vlastně vůbec nepotřebujeme. Přesto definici (3.8) uvádíme, protože jednak jste na ní zvyklí z lineární algebry a jednak funguje i pro zobrazení mezi obecnými (tj. axiomaty definovanými) lineárními prostory.

Důkaz. Zobrazení (3.10) je lineární, protože $\mathbf{f}(\mathbf{x} + \mathbf{y}) = \mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} = \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{y})$ a $\mathbf{f}(\alpha\mathbf{x}) = \mathbf{A}(\alpha\mathbf{x}) = \alpha\mathbf{A}\mathbf{x} = \alpha\mathbf{f}(\mathbf{x})$.

Nechť $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je lineární zobrazení. Necht' $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ je standardní báze prostoru \mathbb{R}^n . Pro každé $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ máme $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$. Z (3.8) plyne

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

Vidíme, že \mathbf{A} je matice se sloupci $\mathbf{f}(\mathbf{e}_1), \dots, \mathbf{f}(\mathbf{e}_n)$.

Matice \mathbf{A} je zobrazením \mathbf{f} určena jednoznačně, protože 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 $\mathbf{e}_1, \dots, \mathbf{e}_n$. ■

Platí-li (3.10), říkáme, že matice \mathbf{A} reprezentuje lineární zobrazení \mathbf{f} .

Pokud $m = 1$, lineární zobrazení je funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a obvykle se mu říká *lineární funkce* nebo *lineární forma*. Pak je zvykem ji psát jako

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = a_1x_1 + \dots + a_nx_n, \quad (3.11)$$

což je skalární součin vektorů $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$ neboli maticový součin matice $\mathbf{A} = \mathbf{a}^T \in \mathbb{R}^{1 \times n}$ a vektoru \mathbf{x} .

Příklad 3.3. 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 (3.8). Ovšem je to patrné na první pohled, protože jej lze vyjádřit ve tvaru (3.10):

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

Výraz $\mathbf{A}\mathbf{x}$ ve vzorci (3.10) je maticový součin matice $m \times n$ s maticí $n \times 1$, tj. se sloupcovým vektorem (viz §2.2). Je důležité se s tímto výrazem beze zbytku seznámit, protože ho v různém značení budeme potkávat znovu a znovu. Označíme-li $\mathbf{y} = \mathbf{A}\mathbf{x}$, podle (2.2) je

$$\begin{aligned} y_1 &= a_{11}x_1 + \dots + a_{1n}x_n, \\ &\vdots \\ y_m &= a_{m1}x_1 + \dots + a_{mn}x_n. \end{aligned} \quad (3.12)$$

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

$$\mathbf{A}\mathbf{x} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n, \quad (3.13)$$

tedy vektor $\mathbf{A}\mathbf{x}$ je *lineární kombinace sloupců* matice \mathbf{A} s koeficienty \mathbf{x} . Tohle si dobře uvědomte! Plyne z toho například, že definice lineární nezávislosti (3.2) pro sloupce matice \mathbf{A} zní

$$\mathbf{A}\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0} \quad (3.14)$$

(nedejte se poplést tím, že \mathbf{x} v (3.14) označuje koeficienty lineární kombinace, zatímco v (3.2) označuje kombinované vektory!).

Napíšeme-li matici \mathbf{A} pomocí řádků, máme

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{x} \\ \vdots \\ \mathbf{a}_m^T \mathbf{x} \end{bmatrix}, \quad (3.15)$$

tedy složky vektoru \mathbf{Ax} jsou *skalární součiny řádků* matice \mathbf{A} a vektoru \mathbf{x} . Všimněte si, že (3.13) a (3.15) jsou speciální případy (2.21) a (2.20).

Zde je snadný ale důležitý fakt o skládání (viz §1.1.2) lineárních zobrazení:

Tvrzení 3.7. Složení lineárních zobrazení je opět lineární zobrazení, přičemž matice složeného zobrazení je součinem matic jednotlivých zobrazení.

Důkaz. Tvrzení stačí dokázat pro dvě zobrazení, pro více plyne z asociativity skládání zobrazení. Je-li $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ a $\mathbf{g}(\mathbf{y}) = \mathbf{By}$, pak $(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{B}(\mathbf{Ax}) = (\mathbf{BA})\mathbf{x} = \mathbf{BAx}$. ■

Všimněte si, že v důkazu jsme použili asociativitu maticového součinu (viz §2.1.2). Protože je skládání zobrazení asociativní, musí být i maticový součin. To je velmi přirozený důvod, proč maticový součin musí být asociativní.

3.2.1 Prostor obrazů a nulový prostor

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

Prostor obrazů (angl. *range*) matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ (příp. zobrazení \mathbf{f}) je množina⁸

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

Z této definice plyne, že prostor obrazů je

- množina $\mathbf{f}(\mathbb{R}^n)$ všech hodnot, jichž může zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ nabýt.
- množina všech vektorů \mathbf{y} , pro které má lineární soustava $\mathbf{Ax} = \mathbf{y}$ řešení.
- lineární obal sloupců matice \mathbf{A} (viz (3.13)). Tedy je to lineární podprostor \mathbb{R}^m .

Nulový prostor (angl. *null space*) matice \mathbf{A} (příp. zobrazení \mathbf{f}) je množina

$$\text{null } \mathbf{A} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{0} \}. \quad (3.17)$$

Z této definice plyne, že nulový prostor je

- množina všech vektorů, které se zobrazí do nulového vektoru.
- množina všech vektorů, které jsou ortogonální na každý řádek matice \mathbf{A} (viz (3.15)). Z toho je vidět, že je to lineární podprostor \mathbb{R}^n .

Pro prostor obrazů a nulový prostor se často používají i jiné názvy. Prostoru $\text{null } \mathbf{A}$ se říká také *jádro* nebo *kernel*, což se častěji používá pro lineární zobrazení mezi obecnými (axiomy definovanými) lineárními prostory. Prostoru $\text{rng } \mathbf{A}$ se říká také *sloupcový prostor* matice \mathbf{A} , protože je to lineární obal sloupců. Lineárnímu obalu řádků $\text{rng}(\mathbf{A}^T)$ se pak říká *řádkový prostor*

⁸Někdo raději píše $\text{rng}(\mathbf{A})$ než $\text{rng } \mathbf{A}$, ale přece je normální psát např. $\sin x$ a ne $\sin(x)$.

matice \mathbf{A} . Podobně, prostorům $\text{null } \mathbf{A}$ resp. $\text{null}(\mathbf{A}^T)$ se přesněji říká *pravý nulový prostor* resp. *levý nulový prostor* matice \mathbf{A} . Čtyřem podprostorům

$$\text{rng } \mathbf{A}, \quad \text{rng}(\mathbf{A}^T), \quad \text{null } \mathbf{A}, \quad \text{null}(\mathbf{A}^T) \quad (3.18)$$

se někdy říká *základní* (nebo *fundamentální*) *podprostory* generované maticí \mathbf{A} . Uvidíme později v §4.5, že tyto podprostory jsou po dvojicích svázány operací ortogonálního doplňku.

3.2.2 Hodnost

Hodnost matice je definována jako dimenze lineárního obalu jejích sloupců,

$$\text{rank } \mathbf{A} = \dim \text{rng } \mathbf{A}. \quad (3.19)$$

Věta 3.8. 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 $\mathbf{A} = \mathbf{BC}$.

Důkaz. Zvolme libovolnou bázi prostoru $\text{rng } \mathbf{A}$. Necht' tato báze tvoří sloupce matice $\mathbf{B} \in \mathbb{R}^{m \times r}$, tedy $r = \text{rank } \mathbf{A} = \dim \text{rng } \mathbf{A}$. Nyní j -tý sloupec \mathbf{a}_j matice \mathbf{A} je lineární kombinací sloupců matice \mathbf{B} , neboli $\mathbf{a}_j = \mathbf{Bc}_j$ pro nějaké $\mathbf{c}_j \in \mathbb{R}^r$. Je tedy $\mathbf{A} = \mathbf{BC}$, kde matice $\mathbf{C} \in \mathbb{R}^{r \times n}$ má sloupce $\mathbf{c}_1, \dots, \mathbf{c}_n$. ■

Rozkladu dle Věty 3.8 se říká **rozklad matice podle hodnosti** (angl. *rank factorization*). Lze jej vidět jako 'kompresi' matice \mathbf{A} , což při $r \ll \min\{m, n\}$ může být značná úspora. Např.:

- Uložení matice \mathbf{A} do paměti zabere mn čísel, uložení matic \mathbf{B} a \mathbf{C} jen $(m+n)r$ čísel.
- Přímý výpočet maticového součinu \mathbf{Ax} vyžaduje $O(mn)$ operací. Spočítáme-li ale nejdřív vektor $\mathbf{z} = \mathbf{Cx}$ a pak vektor $\mathbf{y} = \mathbf{Bz}$, potřebujeme jen $O((m+n)r)$ operací.
- Představme si zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ jako přenosový kanál, jehož vstupem jsou proměnné x_1, \dots, x_n a výstupem y_1, \dots, y_m . Přenos můžeme realizovat jen po r 'drátech' z_1, \dots, z_r .

S pomocí Věty 3.8 dokážeme zásadní fakt, že dimenze lineárního obalu sloupců je rovna dimenzi lineárního obalu řádků:

Věta 3.9. Pro každou matici \mathbf{A} platí $\text{rank } \mathbf{A} = \text{rank}(\mathbf{A}^T)$.

Důkaz. Pišme $\mathbf{A} = \mathbf{BC}$ ve Větě 3.8 jako $\mathbf{A}^T = \mathbf{C}^T \mathbf{B}^T$. Z Tvrzení 3.12 je $\text{rng}(\mathbf{A}^T) \subseteq \text{rng}(\mathbf{C}^T)$. Máme tedy

$$\text{rank}(\mathbf{A}^T) = \dim \text{rng}(\mathbf{A}^T) \leq \dim \text{rng}(\mathbf{C}^T) \leq r = \text{rank } \mathbf{A}, \quad (3.20)$$

kde první nerovnost plyne z Věty 3.3 a druhá nerovnost z toho, že matice \mathbf{C}^T má r sloupců.

Ale nerovnost (3.20) platí pro každou matici \mathbf{A} . Můžeme ji tedy použít na \mathbf{A}^T , čímž dostaneme $\text{rank } \mathbf{A} \leq \text{rank}(\mathbf{A}^T)$. Obě nerovnosti dohromady dávají $\text{rank } \mathbf{A} = \text{rank}(\mathbf{A}^T)$. ■

Z Věty 3.2 plyne, že dimenze lineárního obalu množiny vektorů nemůže být větší než počet těchto vektorů. To spolu s Větou 3.9 znamená, že pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ je

$$\text{rank } \mathbf{A} \leq \min\{m, n\}. \quad (3.21)$$

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.

Z Věty (3.21) plyne, že každá množina více než n vektorů v \mathbb{R}^n je lineárně závislá. Tento známý fakt se může zdát očividný, ale pro jeho důkaz jsme potřebovali (netriviální) Větu 3.9.

Je užitečné si pamatovat horní mez na hodnost součinu matic (důkaz viz Cvičení 3.15):

$$\text{rank}(\mathbf{AB}) \leq \min\{\text{rank } \mathbf{A}, \text{rank } \mathbf{B}\}. \quad (3.22)$$

3.2.3 Matice s plnou hodnotí

Matice s plnou hodnotí, tj. s lineárně nezávislými sloupci nebo lineárně nezávislými řádky, jsou důležité a ‘hezké’. Uvedeme řadu vlastností, které jsou tomu ekvivalentní. Důkazy vět jsou jednoduché, spíše jde o to shrnout a uspořádat dosavadní poznatky (a nezamotat se v tom).

Věta 3.10. Pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou následující výroky 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, tj. $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^m$ (viz §1.1.2).
5. Řádky matice \mathbf{A} jsou lineárně nezávislé.
6. Matice \mathbf{A} má pravou inverzi, tj. $\mathbf{AB} = \mathbf{I}$ pro nějakou matici \mathbf{B} .
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 hodnoti a z Věty 3.9.
- $2 \Rightarrow 6$ 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 (5.5a), uvedené později. ■

Věta 3.11. Pro matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou následující výroky 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. $\mathbf{BA} = \mathbf{I}$ pro nějakou matici \mathbf{B} .
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 (3.17).
- 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 hodnoti (3.19).
- 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 ekvivalence $3 \Leftrightarrow 6$ ve Větě 3.10.
- Tvrzení $1 \Leftrightarrow 7$ je plyne z rovnosti (5.5b), uvedené později. ■

Z těchto vět plyne, že čtvercová matice má plnou hodnost, právě když je regulární (viz §2.1.4). Nakonec uvedeme snadné tvrzení, které se bude později mockrát hodit:

Tvrzení 3.12. *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.*
- $\text{null}(\mathbf{AB}) \supseteq \text{null } \mathbf{B}$.
- $\text{null}(\mathbf{AB}) = \text{null } \mathbf{B}$, *jestliže sloupce matice \mathbf{A} jsou lineárně nezávislé.*

Důkaz.

- Máme dokázat, že jestliže soustava $\mathbf{z} = \mathbf{ABx}$ má řešení, pak soustava $\mathbf{z} = \mathbf{Ay}$ má řešení. To je ale jasné, protože vezmeme $\mathbf{y} = \mathbf{Bx}$.
- Druhé tvrzení navíc říká, že jestliže má \mathbf{A} lineárně nezávislé řádky a soustava $\mathbf{z} = \mathbf{Ay}$ má řešení, pak soustava $\mathbf{z} = \mathbf{ABx}$ má řešení. To platí, neboť dle Věty 3.10 má soustava $\mathbf{Bx} = \mathbf{y}$ řešení pro každé \mathbf{y} .
- Vynásobíme-li rovnost $\mathbf{Bx} = \mathbf{0}$ maticí \mathbf{A} zleva, dostaneme $\mathbf{ABx} = \mathbf{0}$.
- Druhé tvrzení říká, že když \mathbf{A} má lineárně nezávislé sloupce, pak $\mathbf{ABx} = \mathbf{0} \Rightarrow \mathbf{Bx} = \mathbf{0}$. To platí, neboť dle Věty 3.11 platí $\mathbf{Ay} = \mathbf{0} \Rightarrow \mathbf{y} = \mathbf{0}$. ■

3.2.4 Věta o hodnosti a nulitě

Zatímco dimenze prostoru obrazů matice se jinak nazývá hodnost, dimenzi nulového prostoru matice se někdy říká **nulita** matice. Zopakujeme nyní veledůležitou větu o vztahu hodnosti a nulity, známou jako *rank-plus-nullity theorem*.

Věta 3.13. *Pro každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ platí*

$$\underbrace{\dim \text{rng } \mathbf{A}}_{\text{rank } \mathbf{A}} + \dim \text{null } \mathbf{A} = n. \quad (3.23)$$

Důkaz. Necht' báze prostoru $\text{rng } \mathbf{A}$ jsou sloupce matice $\mathbf{B} \in \mathbb{R}^{m \times r}$. Tedy existuje matice $\mathbf{C} \in \mathbb{R}^{n \times r}$ tak, že $\mathbf{B} = \mathbf{AC}$. Necht' báze prostoru $\text{null } \mathbf{A}$ jsou sloupce matice $\mathbf{D} \in \mathbb{R}^{n \times q}$.

Protože sloupce \mathbf{B} jsou báze $\text{rng } \mathbf{A}$, pro každý vektor $\mathbf{x} \in \mathbb{R}^n$ existuje právě jeden vektor $\mathbf{y} \in \mathbb{R}^r$ tak, že $\mathbf{Ax} = \mathbf{By}$, to jest $\mathbf{Ax} = \mathbf{ACy}$, to jest $\mathbf{A}(\mathbf{x} - \mathbf{Cy}) = \mathbf{0}$. To ale znamená $\mathbf{x} - \mathbf{Cy} \in \text{null } \mathbf{A}$, proto existuje právě jeden vektor $\mathbf{z} \in \mathbb{R}^q$ tak, že $\mathbf{x} - \mathbf{Cy} = \mathbf{Dz}$.

A jsme hotovi. Ukázali jsme totiž, že pro každé $\mathbf{x} \in \mathbb{R}^n$ existuje právě jedno $\mathbf{y} \in \mathbb{R}^r$ a právě jedno $\mathbf{z} \in \mathbb{R}^q$ tak, že $\mathbf{x} = \mathbf{Cy} + \mathbf{Dz}$. To ale znamená, že sloupce matice $[\mathbf{C} \ \mathbf{D}] \in \mathbb{R}^{n \times (r+q)}$ jsou báze prostoru \mathbb{R}^n . Tedy musí být $r + q = n$, což je rovnost (3.23). ■

Interpretace věty 3.13:

- Každá dimenze na vstupu zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ se buď 'splácne' do nulového vektoru nebo se objeví na výstupu.
- Počet lineárně nezávislých řešení lineární homogenní soustavy $\mathbf{Ax} = \mathbf{0}$ je $n - \text{rank } \mathbf{A}$.

- Dle Věty 3.9 je rovnost (3.23) ekvivalentní rovnosti

$$\dim \text{rng}(\mathbf{A}^T) + \dim \text{null } \mathbf{A} = n. \quad (3.24)$$

Zde $\text{rng}(\mathbf{A}^T)$ je lineární obal řádků matice \mathbf{A} (řádkový prostor \mathbf{A}) a $\text{null } \mathbf{A}$ je množina všech vektorů ortogonálních na řádky \mathbf{A} . V §4.5 ukážeme, že tyto dva podprostory jsou ortogonální doplněk jeden druhého. Rovnost (3.24) tedy říká, že součet dimenzí podprostoru a jeho ortogonálního doplnku je n (Věta 4.6).

3.2.5 Projektor

Čtvercová matice $\mathbf{P} \in \mathbb{R}^{n \times n}$ se nazývá **projektor**⁹, jestliže

$$\mathbf{P}^2 = \mathbf{P} \quad (3.25)$$

(kde $\mathbf{P}^2 = \mathbf{P}\mathbf{P}$ značí součin matice \mathbf{P} se sebou, viz (2.3)). Takovým maticím se také říká *idempotentní*, protože transformace $\mathbf{f}(\mathbf{y}) = \mathbf{P}\mathbf{y}$ je idempotentní, tzn. $\mathbf{f} \circ \mathbf{f} = \mathbf{f}$ (viz §1.1.2).

Tvrzení 3.14. Pro každý projektor \mathbf{P} platí:

- $\text{null } \mathbf{P} = \text{rng}(\mathbf{I} - \mathbf{P})$
- Pro každé $\mathbf{x} \in \text{rng } \mathbf{P}$ je $\mathbf{P}\mathbf{x} = \mathbf{x}$.
- Podprostory $\text{rng } \mathbf{P}$ a $\text{null } \mathbf{P}$ jsou komplementární (viz §3.1.1).

Důkaz.

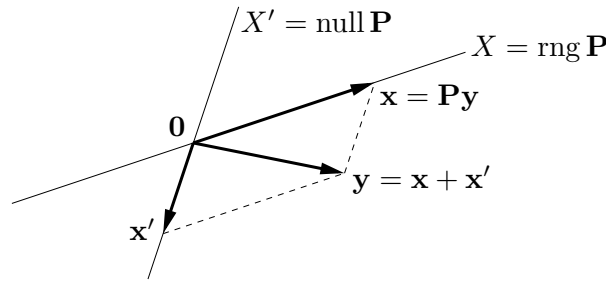
- Necht' $\mathbf{x} \in \text{rng}(\mathbf{I} - \mathbf{P})$, tedy $\mathbf{x} = (\mathbf{I} - \mathbf{P})\mathbf{y}$ pro nějaké $\mathbf{y} \in \mathbb{R}^n$. Pak $\mathbf{P}\mathbf{x} = \mathbf{P}(\mathbf{I} - \mathbf{P})\mathbf{y} = (\mathbf{P} - \mathbf{P}^2)\mathbf{y} = (\mathbf{P} - \mathbf{P})\mathbf{y} = \mathbf{0}$, tedy $\mathbf{x} \in \text{null } \mathbf{P}$.
Necht' $\mathbf{x} \in \text{null } \mathbf{P}$, tedy $\mathbf{P}\mathbf{x} = \mathbf{0}$. Pak $\mathbf{x} = \mathbf{x} - \mathbf{P}\mathbf{x} = (\mathbf{I} - \mathbf{P})\mathbf{x}$, tedy $\mathbf{x} \in \text{rng}(\mathbf{I} - \mathbf{P})$.
- Je $\mathbf{x} \in \text{rng } \mathbf{P}$ právě když $\mathbf{x} = \mathbf{P}\mathbf{y}$ pro nějaké $\mathbf{y} \in \mathbb{R}^n$. Pak $\mathbf{P}\mathbf{x} = \mathbf{P}^2\mathbf{y} = \mathbf{P}\mathbf{y} = \mathbf{x}$.
- Rovnost (3.6b) plyne z Věty 3.13. Rovnost (3.6a) dokážeme takto. Necht' $\mathbf{x} \in \text{rng } \mathbf{P}$, tedy $\mathbf{x} = \mathbf{P}\mathbf{y}$ pro nějaké \mathbf{y} . Jestliže navíc $\mathbf{x} \in \text{null } \mathbf{P}$, je $\mathbf{P}\mathbf{x} = \mathbf{P}\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{y} = \mathbf{x} = \mathbf{0}$. ■

Ve Větě 3.5 jsme ukázali, že máme-li komplementární dvojici podprostorů X, X' , pak každý vektor $\mathbf{y} \in \mathbb{R}^n$ lze jednoznačně rozložit jako $\mathbf{y} = \mathbf{x} + \mathbf{x}'$ kde $\mathbf{x} \in X$ a $\mathbf{x}' \in X'$. Vektor \mathbf{x} jsme nazvali *projekce* vektoru \mathbf{y} na podprostor X ve směru podprostoru X' . Dosazením $\mathbf{x}' = \mathbf{y} - \mathbf{x}$ vidíme, že projekce \mathbf{x} je (jednoznačně) určena podmínkami

$$\mathbf{x} \in X, \quad \mathbf{y} - \mathbf{x} \in X'. \quad (3.26)$$

Tvrzení 3.14 říká, že projektor reprezentuje projekci na $X = \text{rng } \mathbf{P}$ ve směru $X' = \text{null } \mathbf{P}$, neboť vektor $\mathbf{x} = \mathbf{P}\mathbf{y}$ splňuje (3.26). Podmínka idempotence (3.25) také získala zřejmý význam: když nějaký vektor jednou promítneme, další projekce (na stejný podprostor) ho již nezmění. Vše je vidět z obrázku:

⁹Slovo *projektor* má tento význam v lineární algebře. V jiných částech matematiky se toto slovo může používat v trochu jiném významu.



Příklad 3.4. Příkladem projekce v \mathbb{R}^3 je stín vržený na zem předmětem osvětleným sluncem. Podprostor X je rovina země a podprostor X' je přímka rovnoběžná se slunečními paprsky (předpokládáme, že slunce je nekonečně daleko a tedy všechny paprsky jsou rovnoběžné). Když je slunce přímo nad hlavou (paprsky jsou svislé), projekce je ortogonální. V tomto příkladu zanedbáváme zákryty, představte si např., že předmět je mrak izolovaných bodů. ♦

Ukažme, že každá projekce je reprezentována nějakým projektorem, tedy projekce lze charakterizovat¹⁰ jako *lineární idempotentní transformace*:

Věta 3.15. Pro každou dvojici komplementárních podprostorů $X, X' \subseteq \mathbb{R}^n$ existuje právě jeden projektor \mathbf{P} tak, že $X = \text{rng } \mathbf{P}$ a $X' = \text{null } \mathbf{P}$.

Důkaz. Označme $\dim X = \dim \text{rng } \mathbf{P} = \text{rank } \mathbf{P} = m$ a $\dim X' = n - m$. Dle Tvzení 3.4 existuje regulární matice $\mathbf{C} = [\mathbf{B} \ \mathbf{B}'] \in \mathbb{R}^{n \times n}$ tak, že $\mathbf{B} \in \mathbb{R}^{n \times m}$ je báze X a \mathbf{B}' je báze X' . Necht'

$$\mathbf{P} = \mathbf{C} \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{C}^{-1} = \underbrace{\mathbf{C} \begin{bmatrix} \mathbf{I}_m \\ \mathbf{0} \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \mathbf{I}_m & \mathbf{0} \end{bmatrix} \mathbf{C}^{-1}}_{\mathbf{A}}. \quad (3.27)$$

Protože \mathbf{A} má lineárně nezávislé řádky a \mathbf{B} má lineárně nezávislé sloupce, dle Tvzení 3.12 je $\text{rng } \mathbf{P} = \text{rng } \mathbf{B}$ a $\text{null } \mathbf{P} = \text{null } \mathbf{A}$. Protože $X = \text{rng } \mathbf{B}$ a $X' = \text{rng } \mathbf{B}'$, zbývá dokázat $\text{rng } \mathbf{B}' = \text{null } \mathbf{A}$. Je $\mathbf{A}\mathbf{B}' = \mathbf{0}$, tedy $\text{rng } \mathbf{B}' \subseteq \text{null } \mathbf{A}$. Dle Věty 3.13 je $\dim \text{null } \mathbf{A} = n - \text{rank } \mathbf{A} = n - m$. Protože $\dim \text{rng } \mathbf{B}' = n - m$, je $\dim \text{rng } \mathbf{B}' = \dim \text{null } \mathbf{A}$. Dle Tvzení 3.3 tedy $\text{rng } \mathbf{B}' = \text{null } \mathbf{A}$.

Dle Věty 3.5 existuje transformace $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tak, že pro každé \mathbf{y} je $\mathbf{x} = \mathbf{f}(\mathbf{y})$ projekce vektoru \mathbf{y} na X ve směru X' . Ukázali jsme, že tato transformace je reprezentována projektorem \mathbf{P} . Dle Věty 3.6 je matice \mathbf{P} transformací \mathbf{f} určena jednoznačně. ■

Vzorec (3.27) říká, že projektor je podobný¹¹ diagonální matici, která má na diagonále jedničky a nuly (každý projektor je tedy *diagonalizovatelný* a (3.27) je jeho *spektrální rozklad*, jak se dozvíte v §6.1). Tedy ve vhodné bázi (sloupce \mathbf{C}) projektor jen vynuluje některé souřadnice.

Možná vás při čtení §2.1.4 napadla otázka, že když $\mathbf{A}\mathbf{B} = \mathbf{I}$ (tedy \mathbf{A}, \mathbf{B} jsou ve vztahu levá-pravá inverze), co pak asi znamená součin $\mathbf{A}\mathbf{B}$? Zde je odpověď:

Věta 3.16. Matice \mathbf{P} je projektor, právě když $\mathbf{P} = \mathbf{B}\mathbf{A}$ pro nějaké matice \mathbf{A}, \mathbf{B} splňující $\mathbf{A}\mathbf{B} = \mathbf{I}$. V tom případě platí $\text{rng } \mathbf{P} = \text{rng } \mathbf{B}$ a $\text{null } \mathbf{P} = \text{null } \mathbf{A}$.

¹⁰Slovo *charakterizace* v matematice obvykle znamená nutné a postačující podmínky pro nějakou vlastnost.

¹¹Zopakujme z lineární algebry: matice \mathbf{A} a \mathbf{B} jsou si *podobné*, je-li $\mathbf{B} = \mathbf{C}\mathbf{A}\mathbf{C}^{-1}$ pro nějakou regulární \mathbf{C} .

Důkaz. Jestliže $\mathbf{AB} = \mathbf{I}$ a $\mathbf{P} = \mathbf{BA}$, pak $\mathbf{P}^2 = \mathbf{BABA} = \mathbf{BA} = \mathbf{P}$.

Necht' $\mathbf{P}^2 = \mathbf{P}$. Pak dle Tvzení 3.14 jsou podprostory $\text{rng } \mathbf{P}$ a $\text{null } \mathbf{P}$ komplementární, tedy dle Věty 3.15 existují matice \mathbf{A}, \mathbf{B} v (3.27). Platí

$$\mathbf{AB} = [\mathbf{I}_m \quad \mathbf{0}] \mathbf{C} \mathbf{C}^{-1} \begin{bmatrix} \mathbf{I}_m \\ \mathbf{0} \end{bmatrix} = [\mathbf{I}_m \quad \mathbf{0}] \begin{bmatrix} \mathbf{I}_m \\ \mathbf{0} \end{bmatrix} = \mathbf{I}_m. \quad \blacksquare$$

Věta 3.16 navíc ukazuje, že $\boldsymbol{\alpha} = \mathbf{A}\mathbf{y}$ jsou souřadnice projekce $\mathbf{x} = \mathbf{B}\boldsymbol{\alpha}$ v bázi tvořené sloupci matice \mathbf{A} . V důkazu Věty 3.5 je tedy $\mathbf{A} = [\mathbf{x}_1 \cdots \mathbf{x}_m]$ a $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$.

Příklad 3.5. $\mathbf{P} = \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}$ je projektor v rovině \mathbb{R}^2 , který každý bod promítne na podprostor $\text{rng } \mathbf{P} = \text{span}\{(1, 0)\}$ ve směru $\text{null } \mathbf{P} = \text{span}\{(-a, 1)\}$. Pro $a = 0$ je to dokonce ortogonální projektor, který jednoduše vynuluje druhou souřadnici. Máme $m = \text{rank } \mathbf{P} = 1$ a (3.27) zní

$$\underbrace{\begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}}_{\mathbf{P}} = \underbrace{\begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{C}^{-1}} \underbrace{\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}}_{\mathbf{B}} = \underbrace{\begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} 1 & a \end{bmatrix}}_{\mathbf{A}}.$$

Je $\mathbf{AB} = \mathbf{I}_m = 1$ a $\mathbf{BA} = \mathbf{P}$. ◆

Všimněme si nyní i vektoru \mathbf{x}' ve Větě 3.5, tedy projekce vektoru \mathbf{y} na X' ve směru X . Tato projekce je reprezentována maticí $\mathbf{P}' = \mathbf{I} - \mathbf{P}$, neboť $\mathbf{x}' = \mathbf{y} - \mathbf{x} = \mathbf{y} - \mathbf{P}\mathbf{y} = (\mathbf{I} - \mathbf{P})\mathbf{y} = \mathbf{P}'\mathbf{y}$. Protože $\mathbf{P}'^2 = (\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P}) = \mathbf{I} - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{I} - 2\mathbf{P} + \mathbf{P} = \mathbf{I} - \mathbf{P} = \mathbf{P}'$, matice \mathbf{P}' je také projektor. Dvojici projektorů \mathbf{P}, \mathbf{P}' splňující

$$\mathbf{P} + \mathbf{P}' = \mathbf{I} \quad (3.28)$$

nazýváme **komplementární projektory**¹².

Z důkazu Věty 3.15 vysvítá postup 'výroby' komplementárních dvojic projektorů: libovolnou regulární matici \mathbf{C} rozdělíme vodorovně na dva bloky \mathbf{B} a \mathbf{B}' , její inverzi rozdělíme svisle na dva bloky \mathbf{A} a \mathbf{A}' , a napíšeme identity $\mathbf{C}^{-1}\mathbf{C} = \mathbf{I} = \mathbf{C}\mathbf{C}^{-1}$:

$$\begin{bmatrix} \mathbf{AB} & \mathbf{AB}' \\ \mathbf{A'B} & \mathbf{A'B}' \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A} \\ \mathbf{A}' \end{bmatrix}}_{\mathbf{C}^{-1}} \underbrace{\begin{bmatrix} \mathbf{B} & \mathbf{B}' \end{bmatrix}}_{\mathbf{C}} = \mathbf{I} = \underbrace{\begin{bmatrix} \mathbf{B} & \mathbf{B}' \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{A} \\ \mathbf{A}' \end{bmatrix}}_{\mathbf{C}^{-1}} = \underbrace{\mathbf{BA}}_{\mathbf{P}} + \underbrace{\mathbf{B'A}'}_{\mathbf{P}'}. \quad (3.29)$$

Pravá identita je (3.28) a jejím vynásobením vektorem \mathbf{y} dostaneme (3.7). Z levé identity plyne několik věcí, např. že podprostory $\text{rng } \mathbf{B}$ a $\text{rng } \mathbf{B}'$ jsou komplementární. Domyslete zbytek!

3.2.6 Zrcadlení a reflektory

Reflektor je čtvercová matice \mathbf{R} splňující

$$\mathbf{R}^2 = \mathbf{I}. \quad (3.30)$$

Tedy lineární transformace $\mathbf{f}(\mathbf{y}) = \mathbf{R}\mathbf{y}$ je involuce, tj. $\mathbf{f} \circ \mathbf{f}$ je indentické zobrazení (viz §1.1.2). Této transformaci se také říká *reflexe* neboli *zrcadlení*. Tedy zrcadlení je *involutorní lineární transformace*.

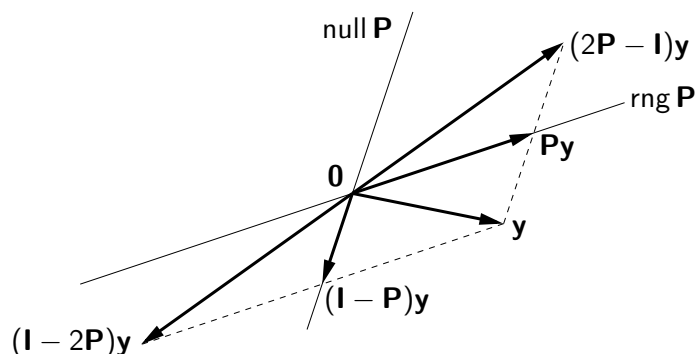
¹²Pozor: zatímco komplementární projektor \mathbf{P}' je svým protějškem \mathbf{P} jednoznačně určen, komplementární podprostor X' není jednoznačně určen podprostorem X (viz Příklad 3.2).

Existuje jednoduchý vztah mezi projektorů a reflektory. Je-li \mathbf{P} projektor, pak

$$\mathbf{R} = 2\mathbf{P} - \mathbf{I} \quad (3.31)$$

je reflektor. Opravdu: $\mathbf{R}^2 = (2\mathbf{P} - \mathbf{I})^2 = 4\mathbf{P}^2 - 2\mathbf{P} - 2\mathbf{P} + \mathbf{I} = 4\mathbf{P} - 2\mathbf{P} - 2\mathbf{P} + \mathbf{I} = \mathbf{I}$ (kde jsme použili idempotenci projektoru). Obráceně, je-li \mathbf{R} reflektor, pak matice $\mathbf{P} = (\mathbf{I} + \mathbf{R})/2$ je idempotentní, tedy projektor (zkontrolujte!).

Z (3.30) navíc vidíme, že je-li \mathbf{R} reflektor, pak $-\mathbf{R}$ je taky reflektor. Reflektor (3.31) zrcadlí každý vektor kolem podprostoru $\text{rng } \mathbf{P}$ ve směru podprostoru $\text{null } \mathbf{P}$, reflektor $-\mathbf{R}$ zrcadlí kolem $\text{null } \mathbf{P}$ ve směru $\text{rng } \mathbf{P}$. To vše je patrné z obrázku, když si promyslíte vztahy mezi nakreslenými vektory. Zároveň vidíme, proč je reflexe involuce: když nějaký vektor ozrcadlíme a výsledný vektor ozrcadlíme znovu (kolem stejného podprostoru a ve stejném směru), pak dostaneme opět ten původní vektor.



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**¹³ lineárního prostoru \mathbb{R}^n , jestliže každá afinní kombinace každé (konečné) množiny vektorů z A leží v A (neboli množina A je uzavřená na afinní kombinace):

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in A, \quad \alpha_1 + \dots + \alpha_k = 1 \quad \implies \quad \alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k \in A. \quad (3.32)$$

Afinní kombinace *nezávisí na počátku*. To znamená, že afinní kombinace vektorů posunutých o libovolný vektor \mathbf{x}_0 je rovna afinní kombinaci neposunutých vektorů posunuté 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, obecná lineární kombinace na počátku závisí.

K tomu (mírně neformální) poznámka: Jestliže s prvky lineárního (pod)prostoru provádíme operace afinní kombinace, tyto prvky si stačí představovat/kreslit jako body. Poloha počátku není důležitá, protože afinní kombinace na něm nezávisí. Afinní kombinaci bodů na papíře lze

¹³Všimněte 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.

sestrojit pomocí pravítka a měřítka bez znalosti polohy počátku. Naproti tomu, jestliže s prvky lineárního (pod)prostoru provádíme operace lineární kombinace, tyto prvky si představujeme jako šipky spojující počátek s koncovým bodem (přesněji jako *volné vektory*), protože lineární kombinaci vektorů můžeme sestrotit pouze se znalostí polohy počátku. To je důvod, proč se prvkům afinního (pod)prostoru říká *body* zatímto prvkům pouhého lineárního (pod)prostoru *vektory*.

Příklad 3.6. Afinní obal dvou bodů $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ je množina

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

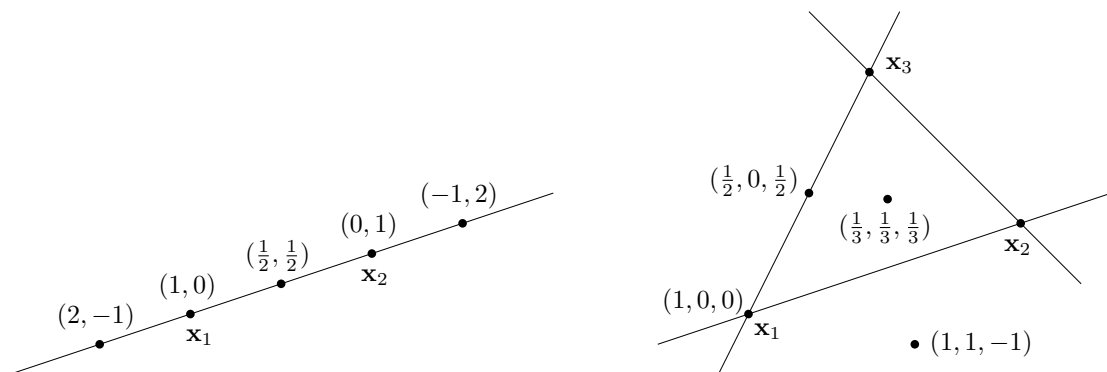
Pokud $\mathbf{x}_1 \neq \mathbf{x}_2$, je touto množinou přímka procházející body $\mathbf{x}_1, \mathbf{x}_2$. Tato přímka je afinní podprostor \mathbb{R}^3 . Na obrázku dole vlevo je několik bodů na přímce a příslušné koeficienty (α_1, α_2) .

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

Afinní obal tří bodů $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in \mathbb{R}^3$ 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}^3 . 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 prostoru \mathbb{R}^3 jsme kreslili jako body, viz poznámka výše.) \blacklozenge

Pro množinu $X \subseteq \mathbb{R}^n$ a vektor $\mathbf{x} \in \mathbb{R}^n$ označme

$$X + \mathbf{x} = \mathbf{x} + X = \{ \mathbf{x} + \mathbf{y} \mid \mathbf{y} \in X \}. \quad (3.33)$$

Věta 3.17.

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

Důkaz. První tvrzení: Chceme dokázat, že afinní kombinace bodů z množiny $X + \mathbf{x}$ leží v této množině. Necht' tedy $\alpha_1 + \dots + \alpha_k = 1$ a chceme dokázat implikaci

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in X + \mathbf{x} \implies \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \in X + \mathbf{x},$$

neboli¹⁴

$$\mathbf{x}_1 - \mathbf{x}, \dots, \mathbf{x}_k - \mathbf{x} \in X \implies \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k - \mathbf{x} = \alpha_1 (\mathbf{x}_1 - \mathbf{x}) + \dots + \alpha_k (\mathbf{x}_k - \mathbf{x}) \in X.$$

Tato implikace platí díky linearitě X .

Druhé tvrzení: Chceme dokázat, že lineární kombinace vektorů z množiny $A - \mathbf{x}$ leží v této množině. Necht' tedy $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ a chceme dokázat implikaci

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in A \implies \alpha_1 (\mathbf{x}_1 - \mathbf{x}) + \dots + \alpha_k (\mathbf{x}_k - \mathbf{x}) \in A - \mathbf{x}.$$

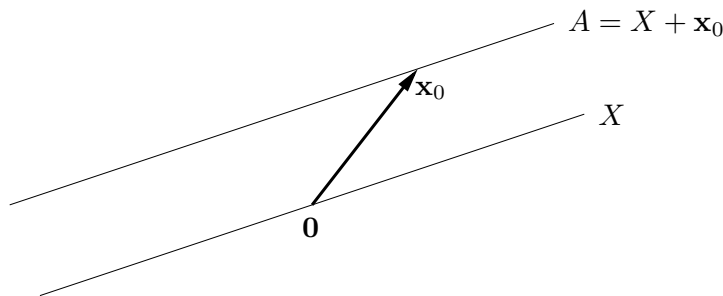
Pravá strana této implikace jde psát jako

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

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

Třetí tvrzení: Označme $X = A - \mathbf{x}$. Jelikož $\mathbf{y} \in A$, je také $\mathbf{y} - \mathbf{x} \in A - \mathbf{x} = X$. Nyní pišme $A - \mathbf{y} = A - \mathbf{x} + \mathbf{x} - \mathbf{y} = X - (\mathbf{y} - \mathbf{x}) = X$, kde poslední rovnost plyne z linearitě X (kterou jsme dokázali v druhém tvrzení). ■

Věta 3.17 říká, že afinní podprostor není nic jiného, než 'posunutý' lineární podprostor (tedy nemusí procházet počátkem, na rozdíl od lineárního podprostoru). Třetí tvrzení věty navíc ukazuje, že tento lineární podprostor je afinním prostorem určen jednoznačně:



Dimenze neprázdného¹⁵ 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**.

Věta 3.18. Množina $A \subseteq \mathbb{R}^n$ je afinní podprostor právě tehdy, když je množinou řešení nějaké lineární soustavy, tj. existují \mathbf{A} a \mathbf{b} splňující $A = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \}$.

¹⁴Zde jsme použili skutečnost, že k výroku typu $\mathbf{x} \in X$ můžeme přičíst libovolný vektor a výrok se tím nezmění. Přesněji, pro každé $X \subseteq \mathbb{R}^n$ a $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ platí $\mathbf{x} \in X \Leftrightarrow \mathbf{x} + \mathbf{y} \in X + \mathbf{y}$. To snadno plyne z (3.33).

¹⁵Na rozdíl od lineárního podprostoru může být afinní podprostor prázdný. Dimenze prázdného afinního podprostoru se někdy konvencí definuje jako -1 .

Důkaz. Předpokládejme, že množina A je neprázdná (pro $A = \emptyset$ věta zjevně platí, protože prázdná množina je afinní podprostor a je to také řešení nějaké lineární soustavy).

Důkaz \Leftarrow : Necht' $A = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}\}$. Necht' $\mathbf{x}_1, \dots, \mathbf{x}_k \in A$ a $\alpha_1 + \dots + \alpha_k = 1$. Dokažme, že $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \in A$:

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

Jiný důkaz \Leftarrow : Necht' $X = \text{null } \mathbf{A}$ a \mathbf{x}_0 je libovolný bod splňující $\mathbf{Ax}_0 = \mathbf{b}$ (tzv. *partikulární řešení* soustavy $\mathbf{Ax} = \mathbf{b}$). Pak

$$A = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}\} = \{\mathbf{x} \mid \mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}\} = \{\mathbf{x}' + \mathbf{x}_0 \mid \mathbf{Ax}' = \mathbf{0}\} = X + \mathbf{x}_0. \quad (3.34)$$

Tedy dle Věty 3.17 je A afinní podprostor.

Důkaz \Rightarrow : Necht' A je afinní podprostor a necht' $\mathbf{x}_0 \in A$. Pak (dle Věty 3.17) je množina $X = A - \mathbf{x}_0$ lineární podprostor, tedy (viz poznámka v §3.2.1) existuje matice \mathbf{A} tak, že $X = \text{null } \mathbf{A}$. Necht' $\mathbf{b} = \mathbf{Ax}_0$. Pak dle (3.34) je $X + \mathbf{x}_0 = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}\}$. ■

Díky Větě 3.18 se afinnímu podprostoru říká také *lineární varieta*. Je to speciální případ *algebraické variety*, což je množina řešení soustavy polynomiálních rovnic. Jejich studiem se zabývá *algebraická geometrie*.

Shrňme: každý lineární podprostor lze reprezentovat buď jako $\text{rng } \mathbf{A}$ pro nějakou matici \mathbf{A} , nebo jako $\text{null } \mathbf{A}$ (tj. množinou řešení soustavy $\mathbf{Ax} = \mathbf{0}$) pro nějakou (jinou!) matici \mathbf{A} . Každý afinní podprostor lze reprezentovat buď jako $\mathbf{x} + X$ pro nějaký vektor \mathbf{x} a lineární podprostor X , nebo jako množinu řešení soustavy $\mathbf{Ax} = \mathbf{b}$ pro nějaké \mathbf{A}, \mathbf{b} .

Body $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ jsou **afinně nezávislé**, jestliže žádný není afinní kombinací ostatních. Afinní nezávislost lze ovšem definovat i jinak (důkaz věty neuvádíme, i když není těžký):

Věta 3.19. Pro body $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ jsou následující výroky ekvivalentní:

1. Žádný bod není roven afinní kombinaci ostatních bodů.
2. Platí (srov. se (3.2))

$$\alpha_1 + \dots + \alpha_k = 0, \quad \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \mathbf{0} \quad \implies \quad \alpha_1 = \dots = \alpha_k = 0. \quad (3.35)$$

3. Vektory $\mathbf{x}_2 - \mathbf{x}_1, \mathbf{x}_3 - \mathbf{x}_1, \dots, \mathbf{x}_k - \mathbf{x}_1$ jsou lineárně nezávislé.
4. Vektory $\begin{bmatrix} \mathbf{x}_1 \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} \mathbf{x}_k \\ 1 \end{bmatrix}$ jsou lineárně nezávislé (kde $\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1}$ značí vektor \mathbf{x} s přidanou jedničkou jako $(n+1)$ -ní souřadnicí¹⁶).

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*). ♦

Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ nazveme **afinní**, pokud (3.8) platí pro všechna $\alpha_1 + \dots + \alpha_k = 1$. 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

$$\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}, \quad (3.36)$$

¹⁶Tento vektor dobře znají počítačové grafiky jako *homogenní souřadnice* bodu \mathbf{x} .

Pro $m = 1$ se zobrazení (3.36) nazývá také **afinní funkce**¹⁷ $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a má tvar

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \cdots + a_n x_n + b, \quad (3.37)$$

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

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 (3.8) pro $\alpha_1 + \cdots + \alpha_k = 1$. Ale je to patrné i z toho, že zobrazení lze vyjádřit ve tvaru (3.36):

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

Příklad 3.9. Často potkáme zobrazení ve tvaru

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}(\mathbf{x} - \mathbf{x}_0). \quad (3.38)$$

To je afinní zobrazení (3.36), kde $\mathbf{b} = -\mathbf{A}\mathbf{x}_0$. Zatímco v (3.36) explicitně vidíme posunutí v hodnotě (vektor \mathbf{b}), v (3.38) explicitně vidíme posunutí v argumentu. \blacklozen

Příklad 3.10. Pro dané body $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ a $\mathbf{y}_1, \dots, \mathbf{y}_k \in \mathbb{R}^m$ máme za úkol najít afinní zobrazení \mathbf{f} takové, aby platilo $\mathbf{y}_i = \mathbf{f}(\mathbf{x}_i)$ pro všechna $i = 1, \dots, k$. Řešíme tedy lineární soustavu $\mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \mathbf{b}$, $i = 1, \dots, k$, pro neznámé $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$. Tuto soustavu můžeme napsat v maticovém tvaru

$$[\mathbf{y}_1 \ \cdots \ \mathbf{y}_k] = [\mathbf{A} \ \mathbf{b}] \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_k \\ 1 & \cdots & 1 \end{bmatrix}.$$

Je-li $k = n + 1$ a body $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou afinně nezávislé, matice $\begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_k \\ 1 & \cdots & 1 \end{bmatrix}$ je regulární a tedy soustava má jediné řešení. \blacklozen

Shrňme: každé lineární zobrazení lze reprezentovat jako $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ pro nějakou \mathbf{A} a každé afinní zobrazení jako $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ pro nějaké \mathbf{A}, \mathbf{b} .

3.4 Cvičení

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

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

¹⁷V lineární algebře znamená ‘lineární funkce’ 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.

e) $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0 \}$

f) $\mathbf{a} + \text{span}\{\mathbf{b}, \mathbf{c}\}$, kde $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ jsou známé vektory takové, že \mathbf{b}, \mathbf{c} jsou lineárně nezávislé.

- 3.2. Je množina $\{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_3 = 0 \}$ lineární podprostor? Pokud ano, najděte jeho libovolnou bázi.
- 3.3. Máme podprostor $X \subseteq \mathbb{R}^3$ s bází $(1, 2, 3), (-1, 0, 1)$. Necht' $\mathbf{x} = (2, 2, 2)$. Platí $\mathbf{x} \in X$? Pokud ano, najděte souřadnice vektoru \mathbf{x} v této bázi.
- 3.4. Dokažte, že jestliže jsou vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ lineárně závislé, pak je aspoň jeden z nich lineární kombinací ostatních.
- 3.5. 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.
- 3.6. 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 dvou vektorů¹⁸. 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} ?
- 3.7. 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 (3.10). Je toto zobrazení afinní? Pokud ano, napište ho ve formě (3.36). Obě odpovědi dokažte z definic.
- 3.8. 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 $\mathbf{x}_0 \in \mathbb{R}^3$.

- 3.9. 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) $f(1, 0, 1) = -1, f(0, 1, 2) = 1, f(1, 1, 3) = 2$.

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

- 3.11. Sloupce matice $\mathbf{A} \in \mathbb{R}^{n \times m}$ jsou báze podprostoru $X \subseteq \mathbb{R}^n$, právě když $\dim X = m$ a $X = \text{rng } \mathbf{A}$. Proč?
- 3.12. Necht' matice \mathbf{A} a \mathbf{B} mají stejné rozměry a obě mají lineárně nezávislé sloupce. Dokažte, že sloupce matic tvoří báze stejného podprostoru, právě když $\mathbf{B} = \mathbf{A}\mathbf{C}$ pro nějakou regulární matici \mathbf{C} .
- 3.13. 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?
- 3.14. Dokažte, že matice $[\mathbf{A} \ \mathbf{B}]$ je regulární, právě když jsou podprostory $\text{rng } \mathbf{A}$ a $\text{rng } \mathbf{B}$ komplementární.

¹⁸Tedy \times zde neoznačuje kartézský součin množin, jako jinde ve skriptech.

- 3.15. Dokažte, že pro libovolné dvě matice platí nerovnost (3.22).
- 3.16. Navrhněte postup, jak spočítat 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.
- 3.17. Které z těchto výroků jsou pravdivé? Každý výrok dokažte nebo najděte protipříklad. Některé výroky mohou platit jen pro určité rozměry matic – najděte co nejobecnější podmínky na rozměry matic, aby výroky byly pravdivé.
- Pokud \mathbf{AB} má plnou hodnost, pak \mathbf{A} a \mathbf{B} mají plnou hodnost.
 - Pokud \mathbf{A} a \mathbf{B} mají plnou hodnost, pak \mathbf{AB} má plnou hodnost.
 - Pokud \mathbf{A} a \mathbf{B} mají triviální nulový prostor, pak \mathbf{AB} má triviální nulový prostor.
 - (\star) Pokud \mathbf{A} a \mathbf{B} jsou úzké s plnou hodností a platí $\mathbf{A}^T \mathbf{B} = \mathbf{0}$, pak matice $[\mathbf{A} \ \mathbf{B}]$ je úzká s plnou hodností.
 - (\star) Pokud matice $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$ má plnou hodnost, pak \mathbf{A} i \mathbf{B} mají plnou hodnost.
- 3.18. Dokažte, že když nějaká matice splňuje $\mathbf{P}^2 = \mathbf{P}$, pak pro každé $\mathbf{x} \in \text{rng } \mathbf{P}$ platí $\mathbf{P}\mathbf{x} = \mathbf{x}$.
- 3.19. Pro projektor dokažte: (a) $\text{rng } \mathbf{P} = \mathbb{R}^n \Rightarrow \mathbf{P} = \mathbf{I}$, (b) $\text{null } \mathbf{P} = \mathbf{0} \Rightarrow \mathbf{P} = \mathbf{I}$.
- 3.20. Je složení dvou projekcí projekce? Odpověď dokažte.
- 3.21. (\star) Mějme vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$. Vektor \mathbf{x}_i nazveme *klíčový*, není-li lineární kombinací vektorů $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}$. Dokažte, že množina klíčových vektorů je báze podprostoru $\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Všimněte si, že vlastně dokazujete první tvrzení Věty 3.2.
- 3.22. (\star) 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 $j \in \{1, \dots, n\} \setminus J$.
- 3.23. (\star) Dokažte Větu 3.19.

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 $n = 1$ a $\mathbf{a} \neq \mathbf{0}$ je množinou jediný bod, tedy afinní podprostor prostoru \mathbb{R} . Pro $n = 1$ a $\mathbf{a} = \mathbf{0}$ je množina prázdná (tedy není afinní podprostor). Pro $n > 1$ je množina také prázdná, protože soustava $\mathbf{a}\mathbf{x}^T = \mathbf{I}$ nemá řešení pro žádné \mathbf{a} , \mathbf{x} (možný důkaz: je $\text{rank } \mathbf{I} = n$, ale $\text{rank}(\mathbf{a}\mathbf{x}^T) \leq 1$).
- 3.1.e) Lineární podprostor dimenze $n - 1$.
- 3.1.f) Je to rovina v \mathbb{R}^n , která neprochází počátkem (prochází bodem \mathbf{a}), tedy je to afinní podprostor dimenze 2.
- 3.2. Ano. Je to vlastně množina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = 0\}$ (tedy nadrovina procházející počátkem s normálovým vektorem \mathbf{a}), kde $n = 4$ a $\mathbf{a} = (1, 0, 1, 0)$. Bázi získáme řešením homogenní soustavy $\mathbf{a}^T \mathbf{x} = x_1 + x_3 = 0$. Báze má tři prvky (protože dimenze nadroviny je $n - 1 = 3$).

- 3.3. Je $\mathbf{x} \in X$, právě když $\mathbf{x} = \alpha\mathbf{a} + \beta\mathbf{b}$ pro nějaké $\alpha, \beta \in \mathbb{R}$. Pokud tato lineární soustava má řešení, pak (α, β) jsou souřadnice vektoru \mathbf{x} v bázi (\mathbf{a}, \mathbf{b}) (zde je drobná nesrovnalost: někdy jsme o bázi mluvili jako o množině vektorů, protože nám na pořadí vektorů báze nezáleželo).
- 3.4. Jsou-li vektory lin. závislé, pak $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k = \mathbf{0}$ kde aspoň jedno z čísel $\alpha_1, \dots, \alpha_k$ je nenulové. Je-li např. $\alpha_1 \neq 0$, pak $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k = \mathbf{0}$ lze napsat jako $\mathbf{x}_1 = \alpha'_2\mathbf{x}_2 + \dots + \alpha'_k\mathbf{x}_k$ kde $\alpha'_i = \alpha_i/\alpha_1$.
- 3.5. Ano. Důkaz: Dle (3.3) máme dokázat, že pro každé $\mathbf{y}_1, \dots, \mathbf{y}_k \in \mathbf{f}(X)$ platí $\sum_i \alpha_i \mathbf{y}_i \in \mathbf{f}(X)$. Protože $\mathbf{f}(X) = \{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in X\}$, je $\mathbf{y} \in \mathbf{f}(X)$ právě když existuje $\mathbf{x} \in X$ tak, že $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Tedy máme dokázat, že pro každé $\mathbf{x}_1, \dots, \mathbf{x}_k \in X$ existuje $\mathbf{x} \in X$ tak, že $\sum_i \alpha_i \mathbf{f}(\mathbf{x}_i) = \mathbf{f}(\mathbf{x})$. To je ale pravda, protože $\sum_i \alpha_i \mathbf{f}(\mathbf{x}_i) = \mathbf{f}(\sum_i \alpha_i \mathbf{x}_i)$ a $\mathbf{x} = \sum_i \alpha_i \mathbf{x}_i \in X$.
- 3.6. 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 pro $\mathbf{y} \neq \mathbf{0}$.
- 3.7. Je afinní. Je $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kde $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{bmatrix}$ $\mathbf{b} = (0, -1, 0) = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{x} = (x, y) = \begin{bmatrix} x \\ y \end{bmatrix}$.
- 3.8. Např. $(1, -1, 2) + \text{span}\{(1, -1, 1)\} = \{(1 + \alpha, -1 - \alpha, 2 + \alpha) \mid \alpha \in \mathbb{R}\}$.
- 3.9.a) Je $f(x_1, x_2) = a_1x_1 + a_2x_2$. Rešíme soustavu $a_1 + 2a_2 = 2$, $3a_1 + 4a_2 = 3$. Tato soustava má řešení, tedy lineární funkce existuje.
- 3.9.b) Ano.
- 3.9.c) Ne.
- 3.10.a) $\text{rng } \mathbf{A} = \mathbb{R}^2$ a tedy báze $\text{rng } \mathbf{A}$ je např. $(0, 1), (1, 0)$. Báze $\text{null } \mathbf{A}$ je např. $(1, 1, 3)$.
- 3.10.b) Báze $\text{rng } \mathbf{A}$ je např. $(2, 1, 1), (1, -1, 2)$. Báze $\text{null } \mathbf{A}$ je např. $(0, 0)$.
- 3.12. Tohle znáte z lineární algebry: \mathbf{C} je matice přechodu mezi bázemi tvořenými sloupci \mathbf{A} a \mathbf{B} .
- 3.13. $\text{rng } \mathbf{A} = \text{rng } \mathbf{B}$ je ekvivalentní $\text{rank } \mathbf{A} = \text{rank}[\mathbf{A} \ \mathbf{B}] = \text{rank } \mathbf{B}$.
- 3.14. Matice je regulární, právě když je čtvercová a má lineárně nezávislé sloupce. Uvědomte si, že lineární kombinaci sloupců matice $[\mathbf{A} \ \mathbf{B}]$ lze psát jako $[\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \mathbf{A}\alpha + \mathbf{B}\beta$ a použijte definici lineární nezávislosti (3.2).
- 3.15. Dle Tvrdění 3.12 je $\text{rng}(\mathbf{A}\mathbf{B}) \subseteq \text{rng } \mathbf{A}$, z toho $\text{rank}(\mathbf{A}\mathbf{B}) \leq \text{rank } \mathbf{A}$. Dle Věty 3.9 je $\text{rank}(\mathbf{A}\mathbf{B}) = \text{rank}((\mathbf{A}\mathbf{B})^T) = \text{rank}(\mathbf{B}^T \mathbf{A}^T) \leq \text{rank}(\mathbf{B}^T) = \text{rank } \mathbf{B}$.
- 3.16. Viz Příklad 3.10.
- 3.20. Ne vždy. Jako důkaz této odpovědi je třeba najít protipříklad, tedy matice \mathbf{P}, \mathbf{Q} takové, že $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{Q}^2 = \mathbf{Q}$ a $(\mathbf{P}\mathbf{Q})^2 \neq \mathbf{P}\mathbf{Q}$.

Kapitola 4

Ortogonalita

4.1 Délky, úhly, vzdálenosti

Na prostoru \mathbb{R}^n přirozeně máme *standardní skalární součin*

$$\mathbf{x}^T \mathbf{y} = x_1 y_1 + \cdots + x_n y_n = \mathbf{y}^T \mathbf{x}. \quad (4.1)$$

Možná jste zvyklí skalární součin značit $\langle \mathbf{x}, \mathbf{y} \rangle$ (nebo $\mathbf{x} \cdot \mathbf{y}$). Toto značení je nutné, když jde o obecný (axiomy definovaný) skalární součin na obecném (axiomy definovaném) lineárním prostoru (který se pak nazývá *lineární prostor se skalárním součinem*)¹. Pro standardní skalární součin na \mathbb{R}^n by to ale byl zbytečný nový symbol a proto ho budeme psát jednoduše jako $\mathbf{x}^T \mathbf{y}$.

Standardní skalární součin indukuje eukleidovskou normu

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}} = (x_1^2 + \cdots + x_n^2)^{1/2}, \quad (4.2)$$

Norma měří *délku* vektoru \mathbf{x} . Splňuje **trojúhelníkovou nerovnost** $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Úhel φ dvojice vektorů se spočítá jako

$$\cos \varphi = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (4.3)$$

Eukleidovská norma indukuje **eukleidovskou metriku**

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|, \quad (4.4)$$

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}^3 se standardním skalárním součinem se často říká *Eukleidovský prostor*.

4.2 Ortogonální a ortonormální vektory

Vektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ jsou navzájem **ortogonální** (někdy také říkáme *kolmé*²) jestliže $\mathbf{x}^T \mathbf{y} = 0$. Což značíme také $\mathbf{x} \perp \mathbf{y}$. Dle (4.3) to znamená, že vektory svírají pravý úhel.

¹Standardní skalární součin na \mathbb{R}^n , definovaný vzorcem (4.1), se anglicky obvykle nazývá *dot product*, zatímco obecný skalární součin na obecném lineárním prostoru je *inner product*.

²*Ortogonální vektory* znamená skoro totéž jako *kolmé vektory* (angl. *perpendicular*). Relace ortogonalit se definuje pro dvojice prvků z libovolného lineárního prostoru (tedy např. pro dvojice funkcí, posloupností, atd.).

Vektor $\mathbf{x} \in \mathbb{R}^n$ nazveme **normalizovaný**, pokud má jednotkovou délku, tj. $\|\mathbf{x}\| = 1 = \mathbf{x}^T \mathbf{x}$. Množinu vektorů $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ nazveme **ortonormální** (trochu nepřesně³ říkáme, že *vektory* $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou ortonormální), jestliže každý vektor z této množiny je normalizovaný a každá dvojice vektorů z této množiny je ortogonální, tedy

$$\mathbf{x}_i^T \mathbf{x}_j = \delta_{ij} = \begin{cases} 0 & \text{když } i \neq j, \\ 1 & \text{když } i = j. \end{cases} \quad (4.5)$$

Věta 4.1. Ortonormální množina vektorů je lineárně nezávislá.

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

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

To platí pro každé i , tedy $\alpha_1 = \dots = \alpha_k = 0$. To je pravá strana implikace (3.2). ■

Věta 4.2. Necht' $\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$, kde vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou ortonormální. Pak

$$\alpha_i = \mathbf{x}_i^T \mathbf{x}, \quad i = 1, \dots, k. \quad (4.6)$$

Důkaz. Vynásobením rovnice $\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ skalárně vektorem \mathbf{x}_i dostaneme $\mathbf{x}_i^T \mathbf{x} = \mathbf{x}_i^T \mathbf{x}_i \alpha_i = \alpha_i$. ■

Skaláry $\alpha_1, \dots, \alpha_k$ jsou souřadnice vektoru \mathbf{x} v ortonormální bázi $\mathbf{x}_1, \dots, \mathbf{x}_k$ podprostoru $\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Vzpomeňte ze střední školy (nakreslete si obrázek!), že výraz (4.6) je délka (se znaménkem) ortogonální projekce vektoru \mathbf{x} do přímky $\text{span}\{\mathbf{x}_i\}$, kde $\|\mathbf{x}_i\| = 1$. 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 koeficienty α_i musíme pracně počítat řešením lineární soustavy.

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

4.3 Matice s ortonormálními sloupci

Necht' sloupce matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ tvoří⁴ ortonormální množinu vektorů. Dle Věty 4.1 jsou sloupce \mathbf{U} lineárně nezávislé, tedy nutně $m \geq n$ (tj. \mathbf{U} je čtvercová nebo úzká). Podmínku ortonormality (4.5) sloupců matice \mathbf{U} lze psát stručně jako (proč?)

$$\mathbf{U}^T \mathbf{U} = \mathbf{I}. \quad (4.7)$$

Relace kolmosti se původně definovala v geometrii pro dvojice *přímek*. Je možné ji rozšířit na dvojice *nenulových* vektorů \mathbf{x}, \mathbf{y} , které vlastně chápeme jako přímky $\text{span}\{\mathbf{x}\}$ a $\text{span}\{\mathbf{y}\}$. Vektory ale musejí být nenulové, neboť $\text{span}\{\mathbf{0}\}$ není přímka. Naproti tomu, nulový vektor $\mathbf{0}$ je ortogonální s každým vektorem, neboť $\mathbf{x}^T \mathbf{0} = 0$ pro každé \mathbf{x} .

³Nepřesně proto, že ortonormalita není vlastnost jednoho vektoru, ale množiny vektorů.

⁴Pro matice s ortonormálními sloupci (§4.3) je obvyklé značit $\mathbf{U}, \mathbf{V}, \mathbf{Q}$ spíše než \mathbf{A}, \mathbf{B} . Podobně, ortonormální vektory se často značí $\mathbf{u}, \mathbf{v}, \mathbf{q}$ a ne $\mathbf{x}, \mathbf{y}, \mathbf{a}$, jak jsme zvyklí.

To znamená, že \mathbf{U}^T je levá inverze matice \mathbf{U} a \mathbf{U} je pravá inverze matice \mathbf{U}^T (viz §2.1.4).

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

$$\mathbf{f}(\mathbf{x})^T \mathbf{f}(\mathbf{y}) = (\mathbf{U}\mathbf{x})^T (\mathbf{U}\mathbf{y}) = \mathbf{x}^T \mathbf{U}^T \mathbf{U} \mathbf{y} = \mathbf{x}^T \mathbf{y}. \quad (4.8)$$

Pro $\mathbf{x} = \mathbf{y}$ dostaneme $\|\mathbf{f}(\mathbf{x})\| = \|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$, neboli zobrazení zachovává také eukleidovskou normu. Zobrazení \mathbf{f} tedy zachovává délky a úhly. Taková zobrazení se nazývají **isometrie**⁵.

Obecněji, skalární součin matic (viz §2.1.7) je také invariantní vůči násobení isometrií zprava i zleva: jsou-li \mathbf{U}, \mathbf{V} matice s ortonormálními sloupci, dle (2.10) máme

$$\langle \mathbf{U}\mathbf{A}\mathbf{V}^T, \mathbf{U}\mathbf{B}\mathbf{V}^T \rangle = \langle \mathbf{A}, \mathbf{U}^T \mathbf{U} \mathbf{B} \mathbf{V}^T \mathbf{V} \rangle = \langle \mathbf{A}, \mathbf{B} \rangle.$$

Tedy i Frobeniova norma (2.11) je invariantní vůči isometrii,

$$\|\mathbf{U}\mathbf{A}\mathbf{V}^T\| = \|\mathbf{A}\|. \quad (4.9)$$

Tvrzení 4.3. Pro každou čtvercovou matici \mathbf{U} platí

$$\mathbf{U}^T \mathbf{U} = \mathbf{I} \iff \mathbf{U}^T = \mathbf{U}^{-1} \iff \mathbf{U}\mathbf{U}^T = \mathbf{I}. \quad (4.10)$$

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. Zbylé implikace dokážeme podobně. ■

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 (4.10) se říká **ortogonální matice**.

Zdůrazněme, že pokud \mathbf{U} je obdélníková s ortonormálními sloupci, neplatí $\mathbf{U}\mathbf{U}^T = \mathbf{I}$. Dále, pokud má \mathbf{U} ortogonální (ne však ortonormální) sloupce, nemusí mít ortogonální řádky⁶.

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 *ortogonální reflexe* neboli zrcadlení (viz §4.6.1).

Příklad 4.1. 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 $\det \mathbf{U} = 1$. ♦

⁵Obecněji, *isometrie* je zobrazení mezi dvěma metrickými (ne nutně lineárními) prostory, které zachovává vzdálenosti. V našem případě bychom mohli přesněji mluvit o *lineární isometrii*, tedy o isometrii, která je zároveň lineární zobrazení.

⁶To 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 4.2. Zrcadlení (neboli reflexe) v \mathbb{R}^2 kolem přímky procházející počátkem a směrnici $\tan(\varphi/2)$ je reprezentováno ortogonální maticí

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

Příklad 4.3. Permutační matice je čtvercová matice, jejíž sloupce jsou permutované vektory standardní báze. Např.

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

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

4.4 Gram-Schmidtova ortonormalizace

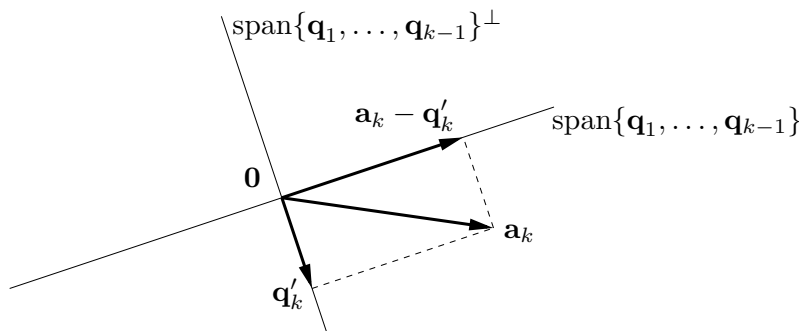
Gram-Schmidtova ortonormalizace je algoritmus, který pro dané lineárně nezávislé vektory $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ najde vektory $\mathbf{q}_1, \dots, \mathbf{q}_n \in \mathbb{R}^m$ takové, že

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

Myšlenka algoritmu je jednoduchá. Předpokládejme, že máme ortonormální vektory $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$ splňující $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{k-1}\}$. Spočítáme vektor

$$\mathbf{q}'_k = \mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{q}_i^T \mathbf{a}_k) \mathbf{q}_i. \quad (4.11)$$

Tento vektor je kolmý na každý z vektorů $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$, protože když skalárně vynásobíme výraz (4.11) jakýmkoliv z nich, tak díky ortonormalitě dostaneme nulu⁷ (ověřte!). Podle předpokladu je $\mathbf{a}_k \notin \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{k-1}\} = \text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$, tedy $\mathbf{q}'_k \neq \mathbf{0}$. Z (4.11) plyne $\mathbf{q}'_k \in \text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}, \mathbf{a}_k\} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$. Teď stačí vektor \mathbf{q}'_k znormalizovat, $\mathbf{q}_k = \mathbf{q}'_k / \|\mathbf{q}'_k\|$.



Algoritmus provede iteraci (4.11) postupně pro $k = 1, \dots, n$. Zde je běh algoritmu pro $n = 3$:

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

⁷Zanedlouho uvidíme, že suma v (4.11) je ortogonální projekce (4.23) vektoru \mathbf{a}_k na ortogonální doplněk podprostoru $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$.

Gram-Schmidtův algoritmem dokazuje nepřekvapivé ale nesamozřejmé vlastnosti ortonormální báze (analogie Věty 3.2):

Věta 4.4 (vlastnosti ortonormální báze).

- Každý lineární podprostor má (alespoň jednu) ortonormální bázi.
- Každou ortonormální množinu vektorů z lineárního podprostoru lze doplnit na ortonormální bázi tohoto podprostoru.

Důkaz. Dle Věty 3.2 má každý podprostor bázi, označme ji $\mathbf{a}_1, \dots, \mathbf{a}_n$. Algoritmus z těchto vektorů vyrobí vektory $\mathbf{q}_1, \dots, \mathbf{q}_n$, které jsou ortonormální báze stejného podprostoru.

Pro druhé tvrzení uvažujme ortonormální množinu vektorů $\mathbf{a}_1, \dots, \mathbf{a}_k$ z nějakého podprostoru. Dle Věty 3.2 tuto množinu doplníme na (ne nutně ortonormální) bázi $\mathbf{a}_1, \dots, \mathbf{a}_n$ tohoto podprostoru, kde $n \geq k$. Algoritmus z vektorů $\mathbf{a}_1, \dots, \mathbf{a}_n$ vyrobí vektory $\mathbf{q}_1, \dots, \mathbf{q}_n$, kde ale z definice algoritmu bude $\mathbf{q}_i = \mathbf{a}_i$ pro každé $i \leq k$. ■

4.4.1 QR rozklad

Věta 4.5 (QR rozklad). Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ lze rozložit na součin

$$\mathbf{A} = \mathbf{QR}, \tag{4.12}$$

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

Rozkladu matice (4.12) se říká **QR rozklad** matice \mathbf{A} . Gram-Schmidtova ortogonalizace umožňuje dokázat Větu 4.5 pro případ, kdy matice \mathbf{A} má lineárně nezávislé sloupce:

Důkaz. Rovnosti (4.11) lze napsat jako

$$\mathbf{a}_k = \sum_{i=1}^{k-1} r_{ik} \mathbf{q}_i + r_{kk} \mathbf{q}_k = \sum_{i=1}^k r_{ik} \mathbf{q}_i, \tag{4.13}$$

kde $r_{ik} = \mathbf{q}_i^T \mathbf{a}_k$ pro $i < k$ a $r_{kk} = \|\mathbf{q}'_k\|$. Soustava rovnic (4.13) pro $k = 1, \dots, n$ se dá napsat v maticovém tvaru jako $\mathbf{A} = \mathbf{QR}$, 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!). ■

Gram-Schmidtův algoritmus je silný teoretický nástroj, ale v uvedené jednoduché podobě je nestabilní vůči zaokrouhlovacím chybám. Jeho modifikacemi lze tuto nevýhodu odstranit a dále ho zobecnit na počítání QR rozkladu libovolné matice. Častěji se ovšem na QR rozklad používají algoritmy založené na *Householderových reflexích* nebo *Givensových rotacích* (neuvádíme).

Rozlišujeme dvě verze QR rozkladu:

- Věta 4.5 popisuje *plnou* verzi QR rozkladu.
- Pokud $m > n$, je posledních $m - n$ řádků matice \mathbf{R} nulových (protože \mathbf{R} je horní trojúhelníková). Tyto řádky jsou násobeny posledními $m - n$ sloupci matice \mathbf{Q} . Můžeme tedy vynechat z matice \mathbf{R} posledních $m - n$ řádků a z matice \mathbf{Q} posledních $m - n$ sloupců, tedy bude $\mathbf{Q} \in \mathbb{R}^{m \times n}$ a $\mathbf{R} \in \mathbb{R}^{n \times n}$. Toto je *redukovaná* verze QR rozkladu.

Pro $m \leq n$ obě verze splývají. V Matlabu spočítáte plný QR rozklad příkazem `[Q,R]=qr(A)` a redukovaný příkazem `[Q,R]=qr(A,0)`. Zkoumejte v Matlabu příkaz `help qr!`

Příklad 4.4. Zde je příklad plného a redukovaného QR rozkladu matice 3×2 :

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{2/3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{2/3} \\ 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} \end{bmatrix}. \quad \blacklozenge$$

QR rozklad je velmi užitečný. Typické je jeho užití na řešení lineárních soustav. Řešíme-li soustavu $\mathbf{Ax} = \mathbf{b}$, rozložíme $\mathbf{A} = \mathbf{QR}$ a vynásobíme soustavu zleva \mathbf{Q}^T , což dá

$$\mathbf{Rx} = \mathbf{Q}^T \mathbf{b}. \quad (4.14)$$

Toto je ekvivalentní úprava, neboť \mathbf{Q} je regulární (v případě plného QR rozkladu). Ale protože je \mathbf{R} trojúhelníková, soustavu jsme velmi zjednodušili. Např. pokud je \mathbf{A} čtvercová regulární, jediné řešení soustavy (4.14) lze levně najít zpětnou substitucí.

4.5 Ortogonální podprostory

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

$$\mathbf{y} \perp \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \iff \mathbf{y} \perp \mathbf{x}_1, \dots, \mathbf{y} \perp \mathbf{x}_k. \quad (4.15)$$

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

$$X \perp Y \implies X \cap Y = \{\mathbf{0}\}, \quad (4.16)$$

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

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

$$X^\perp = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} \perp X\} \quad (4.17)$$

všech vektorů z \mathbb{R}^n ortogonálních na podprostor X . Rozdíl mezi tvrzeními $Y \perp X$ a $Y = X^\perp$ je tedy v tom, že v prvním případě Y nemusí obsahovat *všechny* vektory kolmé na X a ve druhém případě ano. Množina (4.17) je podprostor \mathbb{R}^n (dokažte!).

Příklad 4.5. Následující příklady jsou v prostoru \mathbb{R}^3 . Dvě na sebe kolmé přímky procházející počátkem jsou ortogonální podprostory, jedna přímka není ale ortogonálním doplňkem druhé přímky. Ortogonální doplněk k přímce procházející počátkem je rovina procházející počátkem, která je na tuto přímku kolmá. Ortogonální doplněk celého \mathbb{R}^3 (míněného jako podprostor prostoru \mathbb{R}^3) je počátek $\mathbf{0}$. \blacklozenge

Příklad 4.6. Pozor, stěna místnosti není ortogonální na podlahu. Opravdu, existuje dvojice vektorů, jeden v podlaze a jeden ve stěně, které nejsou ortogonální (kde jsou?)⁸. \blacklozenge

⁸Kolmost se v elementární geometrii definuje pro dvojice přímek, ale tento pojem se rozšiřuje i na dvojice přímka-rovina či rovina-rovina. V tomto pojetí stěna a podlaha na sebe jsou kolmé (s čímž by jistě souhlasil i např. pan zedník). Vidíme tedy, že prostorech \mathbb{R}^n relace kolmosti a ortogonalitě splývají pro dvojice podprostorů dimenze 1 (tj. přímek), ale liší se pro dvojice podprostorů jiných dimenzí.

Věta 4.6. Pro každý podprostor $X \subseteq \mathbb{R}^n$ platí $\dim X + \dim(X^\perp) = n$.

Důkaz. S ortogonálním doplňkem jsme se vlastně už setkali v definici nulového prostoru: nulový prostor matice je tvořen vektory ortogonálními na všechny řádky matice. Dle (4.15) je vektor ortogonální na řádky matice právě tehdy, když je ortogonální na lineární obal řádků matice. Tedy pro každou $\mathbf{A} \in \mathbb{R}^{m \times n}$ je

$$(\text{rng}(\mathbf{A}^T))^\perp = \text{null } \mathbf{A}. \quad (4.18)$$

Dokazovaná rovnost je tedy rovnost (3.24) pro $X = \text{rng}(\mathbf{A}^T)$. ■

Protože $X^\perp \perp X$, dle (4.16) máme $X \cap X^\perp = \{\mathbf{0}\}$. Z Věty 4.6 tedy plyne, že každý podprostor je komplementární se svým ortogonálním doplňkem (viz (3.6)).

Věta 4.7. Pro každý podprostor $X \subseteq \mathbb{R}^n$ platí $(X^\perp)^\perp = X$.

Důkaz. Protože každý vektor je ortogonální na všechny vektory k němu ortogonální, platí $X \subseteq (X^\perp)^\perp$ (tuto větu rozmyslete, opravdu ji chápete?). Z Věty 4.6 použité jednou na X a jednou na X^\perp máme

$$\dim X + \dim X^\perp = n = \dim X^\perp + \dim(X^\perp)^\perp.$$

Z toho plyne $\dim X = \dim(X^\perp)^\perp$. Z Věty 3.3 proto máme $X = (X^\perp)^\perp$. ■

Věta 4.7 ukazuje, že relace ‘býti ortogonálním doplňkem’ je symetrická: $Y = X^\perp \Leftrightarrow X = Y^\perp$. Proto můžeme říkat, že podprostory X a Y jsou *ortogonálním doplňkem jeden druhého*.

4.5.1 Vztah k prostoru obrazů a nulovému prostoru

Rozviňme naše pozorování (4.18) z důkazu Věty 4.6. Vzpomeňme z §3.2.1, že každá matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ definuje čtyři základní podprostory:

- $\text{rng } \mathbf{A} = \{ \mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \} \subseteq \mathbb{R}^m$ je lineární obal sloupců \mathbf{A} ,
- $\text{null } \mathbf{A} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0} \} \subseteq \mathbb{R}^n$ 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} .

Věta 4.8. Pro každou matici \mathbf{A} platí

$$(\text{rng } \mathbf{A})^\perp = \text{null}(\mathbf{A}^T), \quad (4.19a)$$

$$(\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T). \quad (4.19b)$$

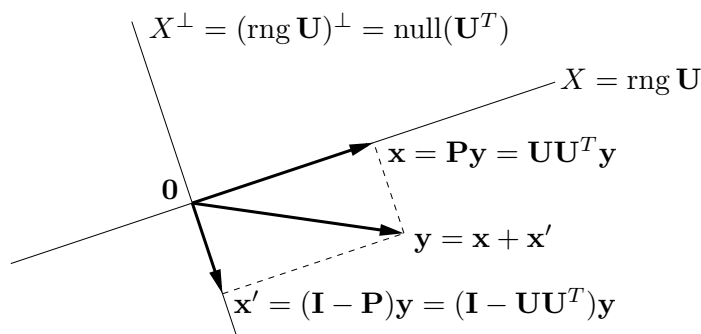
Důkaz. Rovnost (4.19a) plyne z prvního a čtvrtého řádku, je to vlastně rovnost (4.18) použitá na matici \mathbf{A}^T . Rovnost (4.19b) získáme použitím (4.19a) na matici \mathbf{A}^T a rovnosti $(X^\perp)^\perp = X$. ■

4.6 Ortogonální projekce a projektory

V §3.1.1 (zopakujte!) jsme definovali projekci na podprostor ve směru komplementárního podprostoru. Dle (4.16) a Věty 4.6 je každý podprostor komplementární se svým ortogonálním doplňkem. Dle Věty 3.5 lze tedy každý vektor \mathbf{y} jednoznačně rozložit jako $\mathbf{y} = \mathbf{x} + \mathbf{x}'$ kde $\mathbf{x} \in X$ a $\mathbf{x}' \in X^\perp$. Vektoru \mathbf{x} říkáme **ortogonální projekce** vektoru \mathbf{y} na podprostor X (tzn. je to projekce na podprostor X ve směru podprostoru X^\perp). Protože $\mathbf{x}' \in X^\perp \Leftrightarrow \mathbf{x}' \perp X$, vektor \mathbf{x} je jednoznačně určen podmínkami

$$\mathbf{x} \in X, \quad (\mathbf{y} - \mathbf{x}) \perp X. \quad (4.20)$$

Projektor \mathbf{P} reprezentující ortogonální projekci vztahem $\mathbf{x} = \mathbf{P}\mathbf{y}$ (zopakujte §3.2.5!) se nazývá **ortogonální projektor**. Symetricky, \mathbf{x}' je ortogonální projekce vektoru \mathbf{y} na X^\perp a $\mathbf{P}' = \mathbf{I} - \mathbf{P}$ je odpovídající ortogonální projektor (komplementární s \mathbf{P}).



Necht' je podprostor X reprezentován ortonormální bází tvořící sloupce matice $\mathbf{U} \in \mathbb{R}^{n \times m}$, tedy $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ a $X = \text{rng } \mathbf{U}$. Když ve Větě 3.16 položíme $\mathbf{A} = \mathbf{U}^T$ a $\mathbf{B} = \mathbf{U}$, dostaneme vzorec

$$\mathbf{P} = \mathbf{U}\mathbf{U}^T \quad (4.21)$$

pro ortogonální projektor na X . Pozor: ortogonální projektor na podprostor zadaný *obecnou* bází odvodíme až později v §5.1.1.

Pokud jste snad plně nevstřebali látku z §3.1.1 a §3.2.5, raději uvedeme i nezávislé odvození vzorce (4.21) z definice (4.20). Je $\mathbf{x} \in X$ právě když $\mathbf{x} = \mathbf{U}\boldsymbol{\alpha}$, kde $\boldsymbol{\alpha} \in \mathbb{R}^m$ jsou souřadnice vektoru \mathbf{x} v bázi tvořené sloupci \mathbf{U} . Podmínka $(\mathbf{y} - \mathbf{x}) \perp X$ znamená $\mathbf{y} - \mathbf{x} \in X^\perp = \text{null}(\mathbf{U}^T)$, tedy $\mathbf{U}^T(\mathbf{y} - \mathbf{x}) = \mathbf{U}^T(\mathbf{y} - \mathbf{U}\boldsymbol{\alpha}) = \mathbf{0}$. Z toho $\boldsymbol{\alpha} = \mathbf{U}^T \mathbf{y}$. Po dosazení máme $\mathbf{x} = \mathbf{U}\mathbf{U}^T \mathbf{y} = \mathbf{P}\mathbf{y}$.

Označíme-li sloupce matice \mathbf{U} jako $\mathbf{u}_1, \dots, \mathbf{u}_m \in \mathbb{R}^n$, dostaneme vzoreček (4.21) ve formě součtu dyád

$$\mathbf{U}\mathbf{U}^T = [\mathbf{u}_1 \cdots \mathbf{u}_m] [\mathbf{u}_1 \cdots \mathbf{u}_m]^T = \mathbf{u}_1 \mathbf{u}_1^T + \cdots + \mathbf{u}_m \mathbf{u}_m^T. \quad (4.22)$$

Ortogonální projekce vektoru \mathbf{y} na podprostor $X = \text{rng}\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ tedy je

$$\mathbf{x} = \mathbf{U}\mathbf{U}^T \mathbf{y} = \mathbf{u}_1 \mathbf{u}_1^T \mathbf{y} + \cdots + \mathbf{u}_m \mathbf{u}_m^T \mathbf{y} = (\mathbf{u}_1^T \mathbf{y}) \mathbf{u}_1 + \cdots + (\mathbf{u}_m^T \mathbf{y}) \mathbf{u}_m = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_m \mathbf{u}_m, \quad (4.23)$$

kde⁹ $\alpha_i = \mathbf{u}_i^T \mathbf{y}$ jsou jednotlivé souřadnice¹⁰ (viz Věta 4.2). Všimněte si, že každý vektor $\mathbf{u}_i \mathbf{u}_i^T \mathbf{y} = (\mathbf{u}_i^T \mathbf{y}) \mathbf{u}_i = \alpha_i \mathbf{u}_i$ je sám o sobě ortogonální projekce vektoru \mathbf{y} na přímku $\text{rng } \mathbf{u}_i = \text{span}\{\mathbf{u}_i\}$ (kde $\|\mathbf{u}_i\| = 1$) a α_i je délka (se znaménkem) průmětu.

⁹Pozor: závorka ve výrazu $(\mathbf{u}^T \mathbf{y}) \mathbf{u}$ je nutná, protože součin výrazů $\mathbf{u}^T \mathbf{y}$ a \mathbf{u} není maticový součin, ale násobení vektoru skalárem (viz poznámka v §2.1.2). Oproti tomu, výraz $\mathbf{u} \mathbf{u}^T \mathbf{y}$ je maticový součin tří matic $\mathbf{u}, \mathbf{u}^T, \mathbf{y}$, takže závorky v něm být nemusí.

¹⁰Někteří studenti si myslí, že když bod $\mathbf{y} \in \mathbb{R}^n$ promítneme na podprostor menší dimenze $m < n$, tak výsledný vektor $\mathbf{x} = \mathbf{U}\mathbf{U}^T \mathbf{y}$ musí mít m složek. Tak to ale není, protože \mathbf{x} má n složek jako \mathbf{y} . To, co mají tito studenti intuitivně na mysli, je právě vektor souřadnic $\boldsymbol{\alpha} = \mathbf{U}^T \mathbf{y} \in \mathbb{R}^m$.

Následující Věta 4.9 charakterizuje ortogonální projektory: v kombinaci s (3.25) říká, že matice \mathbf{P} je ortogonální projektor právě když

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^T. \quad (4.24)$$

Věta 4.9. Necht' $\mathbf{P}^2 = \mathbf{P}$. Pak $\text{rng } \mathbf{P} \perp \text{null } \mathbf{P}$ platí právě když $\mathbf{P}^T = \mathbf{P}$.

Důkaz. Z $\mathbf{P}^T = \mathbf{P}$ plyne $\text{null } \mathbf{P} = \text{null}(\mathbf{P}^T) = (\text{rng } \mathbf{P})^\perp$, tedy $\text{null } \mathbf{P} \perp \text{rng } \mathbf{P}$.

Tvrzení $\text{rng } \mathbf{P} \perp \text{null } \mathbf{P}$ platí právě když $\text{rng } \mathbf{P} \subseteq (\text{null } \mathbf{P})^\perp = \text{rng}(\mathbf{P}^T) = \text{null}(\mathbf{I} - \mathbf{P}^T)$, kde poslední rovnost plyne z Věty 3.14. To ale znamená $\mathbf{P}(\mathbf{I} - \mathbf{P}^T) = \mathbf{0}$, tj. $\mathbf{P} = \mathbf{P}\mathbf{P}^T$, což po transpozici dá $\mathbf{P}^T = \mathbf{P}\mathbf{P}^T$. Z posledních dvou rovností máme $\mathbf{P} = \mathbf{P}^T$. ■

Nakonec uvedeme snadné ale užitečné tvrzení:

Tvrzení 4.10. Necht' $[\mathbf{U} \ \mathbf{U}']$ je ortogonální matice rozdělená do dvou bloků \mathbf{U}, \mathbf{U}' . Pak:

- Podprostory $\text{rng } \mathbf{U}$ a $\text{rng } \mathbf{U}'$ jsou ortogonálním doplňkem jeden druhého.
- $\mathbf{P} = \mathbf{U}\mathbf{U}^T$ a $\mathbf{P}' = \mathbf{U}'\mathbf{U}'^T$ je komplementární dvojice ortogonálních projektorů (viz (3.28)).

Důkaz. Důkaz kopíruje (3.29): dle Věty 4.3 máme

$$\begin{bmatrix} \mathbf{U}^T\mathbf{U} & \mathbf{U}^T\mathbf{U}' \\ \mathbf{U}'^T\mathbf{U} & \mathbf{U}'^T\mathbf{U}' \end{bmatrix} = \begin{bmatrix} \mathbf{U}^T \\ \mathbf{U}'^T \end{bmatrix} [\mathbf{U} \ \mathbf{U}'] = \mathbf{I} = [\mathbf{U} \ \mathbf{U}'] \begin{bmatrix} \mathbf{U}^T \\ \mathbf{U}'^T \end{bmatrix} = \mathbf{U}\mathbf{U}^T + \mathbf{U}'\mathbf{U}'^T. \quad (4.25)$$

Porovnáním bloků matice vlevo s bloky jednotkové matice \mathbf{I} máme $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{U}'^T\mathbf{U}' = \mathbf{I}$ a $\mathbf{U}^T\mathbf{U}' = \mathbf{0}$. Z prvních dvou rovností plyne, že $\mathbf{P} = \mathbf{U}\mathbf{U}^T$ a $\mathbf{P}' = \mathbf{U}'\mathbf{U}'^T$ jsou projektory. Třetí rovnost znamená $\text{rng } \mathbf{U} \perp \text{rng } \mathbf{U}'$, neboli $\text{rng } \mathbf{U}' \subseteq (\text{rng } \mathbf{U})^\perp$. Z rozměrů matic a Věty 4.6 je ale $\dim \text{rng } \mathbf{U}' = \dim(\text{rng } \mathbf{U})^\perp$, tedy dle Tvrzení 3.3 platí $\text{rng } \mathbf{U}' = (\text{rng } \mathbf{U})^\perp$. ■

4.6.1 Ortogonální zrcadlení a reflektory

Ke každému ortogonálnímu projektoru máme **ortogonální reflektor** ve smyslu §3.2.6. Je-li matice \mathbf{P} symetrická, matice (3.31) je také symetrická (ověřte!). Tedy ortogonální reflektory jsou charakterizované rovnostmi $\mathbf{R}^2 = \mathbf{I}$ a $\mathbf{R}^T = \mathbf{R}$. Z toho plyne $\mathbf{R}^T\mathbf{R} = \mathbf{I}$, takže každý ortogonální reflektor je ortogonální matice (naopak to samozřejmě nemusí platit). Protože $\mathbf{R}^T = \mathbf{R}$ a $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ implikují $\mathbf{R}^2 = \mathbf{I}$, ortogonální reflektory jsou právě symetrické ortogonální matice.

Nejnámější je případ, kdy ortogonálně zrcadlíme kolem nadroviny. Má-li tato nadrovina normálový vektor \mathbf{u} s jednotkovou normou, pak projekce na nadrovinu je $\mathbf{P} = \mathbf{I} - \mathbf{u}\mathbf{u}^T$ a reflektor kolem nadroviny je tedy $\mathbf{R} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$. Tato matice je známa jako *elementární reflektor* neboli *Householderova matice*.

4.6.2 Vzdálenost bodu od podprostoru

Vzdálenost bodu $\mathbf{y} \in \mathbb{R}^m$ od podprostoru $X \subseteq \mathbb{R}^m$ je optimální hodnota úlohy

$$\min_{\mathbf{x} \in X} \|\mathbf{y} - \mathbf{x}\|. \quad (4.26)$$

Intuitivně je zřejmé (viz obrázek na začátku §4.6), že tato úloha nabývá optima v jediném bodě a tento bod je ortogonální projekce bodu \mathbf{y} na podprostor X . Tato vlastnost ortogonální projekce je důležitá pro optimalizaci. Dokažme ji přesně:

Věta 4.11. Pro podprostor $X \subseteq \mathbb{R}^m$ a body $\mathbf{y} \in \mathbb{R}^m$ a $\mathbf{x} \in X$ jsou tato tvrzení ekvivalentní:

1. Bod \mathbf{x} je ortogonální projekcí bodu \mathbf{y} na podprostor X , neboli $(\mathbf{y} - \mathbf{x}) \perp X$.
2. Bod \mathbf{x} je nejbližší bodu \mathbf{y} mezi všemi body podprostoru X .

Důkaz. Pro pohodlí posuneme podprostor X a body \mathbf{x} a \mathbf{y} o vektor $-\mathbf{x}$, čímž se nezmění podprostor X (protože $\mathbf{x} \in X$) ani vzdálenosti bodů a ortogonalita vektorů. Větu tedy stačí dokázat pro $\mathbf{x} = \mathbf{0}$.

Důkaz 1 \Rightarrow 2: Necht' $\mathbf{0}$ je ortogonální projekce bodu \mathbf{y} na podprostor X , tj. $\mathbf{y} \perp X$. Vezměme libovolné $\mathbf{x}' \in X$. Protože $\mathbf{y} \perp X$, je $\mathbf{y} \perp \mathbf{x}'$. Tedy vektory \mathbf{y} a \mathbf{x}' jsou odvěsný pravoúhlého trojúhelníka s přeponou $\mathbf{y} - \mathbf{x}'$. Z Pythagorovy věty (viz Cvičení 4.5) máme

$$\|\mathbf{y} - \mathbf{x}'\|^2 = \|\mathbf{y}\|^2 + \|\mathbf{x}'\|^2 \geq \|\mathbf{y}\|^2.$$

Tedy počátek $\mathbf{0}$ je nejbližší bodu \mathbf{y} mezi všemi body $z \in X$.

Důkaz $\neg 1 \Rightarrow \neg 2$: Necht' $\mathbf{0}$ není ortogonální projekcí bodu \mathbf{y} na podprostor X , tj. $\mathbf{y} \not\perp X$. Tedy existuje $\mathbf{x}' \in X$ tak, že $\mathbf{x}'^T \mathbf{y} \neq 0$. Tedy existuje $\mathbf{x}' \in X$ takové, že $\mathbf{x}'^T \mathbf{x}' = 1$ a $\mathbf{x}'^T \mathbf{y} = \alpha > 0$ (protože vektor \mathbf{x}' vždy můžeme znormalizovat, příp. vynásobit mínus jedničkou). Tedy

$$\|\mathbf{y} - \alpha \mathbf{x}'\|^2 - \|\mathbf{y}\|^2 = \mathbf{y}^T \mathbf{y} - 2\alpha \mathbf{x}'^T \mathbf{y} + \alpha^2 \mathbf{x}'^T \mathbf{x}' - \mathbf{y}^T \mathbf{y} = -2\alpha^2 + \alpha^2 \mathbf{x}'^T \mathbf{x}' = -\alpha^2 < 0.$$

To ale ukazuje, že bod $\alpha \mathbf{x}' \in X$ je bližší bodu \mathbf{y} než počátek. (Všimněte si, že bod $\alpha \mathbf{x}'$ je ortogonální projekce vektoru \mathbf{y} na přímku $\text{span}\{\mathbf{x}'\}$.) ■

Dle Věty 4.11 je vzdálenost bodu \mathbf{y} od podprostoru X rovna délce vektoru \mathbf{x}' a vzdálenost bodu \mathbf{y} od podprostoru X^\perp je délka vektoru \mathbf{x} , kde \mathbf{x} příp. \mathbf{x}' je ortogonální projekce bodu \mathbf{y} na X příp. X^\perp (jako na obrázku výše). Tvoří-li sloupečky matice \mathbf{U} ortonormální bázi podprostoru X , tak délka vektoru \mathbf{x} je

$$\|\mathbf{x}\| = \|\mathbf{U}\mathbf{U}^T \mathbf{y}\| = \|\mathbf{U}^T \mathbf{y}\|, \quad (4.27)$$

kde druhá rovnost platí, protože zobrazení reprezentované maticí \mathbf{U} zachovává eukleidovskou normu (viz §4.3). Délka vektoru \mathbf{x}' je $\|\mathbf{x}'\| = \|(\mathbf{I} - \mathbf{U}\mathbf{U}^T)\mathbf{y}\|$, což už ale nejde zjednodušit. Přidáme-li do hry i matici \mathbf{U}' z Tvrzení 4.10, tvořící ortonormální bázi podprostoru X^\perp , máme

$$\|\mathbf{y}\|^2 = \underbrace{\|\mathbf{U}^T \mathbf{y}\|}_{\|\mathbf{x}\|}^2 + \underbrace{\|\mathbf{U}'^T \mathbf{y}\|}_{\|\mathbf{x}'\|}^2, \quad (4.28)$$

což není nic jiného než Pythagorova věta (viz Cvičení 4.5).

Vzorec (4.27) počítá vzdálenost bodu od lineárního podprostoru $X^\perp = \text{null}(\mathbf{U}^T)$, tj. tento podprostor je tvořen řešeními homogenní soustavy $\mathbf{U}^T \mathbf{x} = \mathbf{0}$. Často potřebujeme spočítat vzdálenost bodu od *afinního* podprostoru. Jak to uděláme? Necht' tento afinní podprostor, $A \subseteq \mathbb{R}^m$, je množinou řešení lineární soustavy $\mathbf{U}^T \mathbf{x} = \mathbf{b}$ (viz Věta 3.18), tedy

$$A = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{U}^T \mathbf{x} = \mathbf{b}\} = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{U}^T(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}\} = X^\perp + \mathbf{x}_0, \quad (4.29)$$

kde \mathbf{x}_0 je libovolný bod splňující $\mathbf{U}^T \mathbf{x}_0 = \mathbf{b}$. Vidíme, že vzdálenost bodu \mathbf{y} od afinního podprostoru A je rovna vzdálenosti bodu $\mathbf{y} - \mathbf{x}_0$ od lineárního podprostoru X^\perp , což je

$$\|\mathbf{U}^T(\mathbf{y} - \mathbf{x}_0)\| = \|\mathbf{U}^T \mathbf{y} - \mathbf{b}\|. \quad (4.30)$$

Speciálně, necht' afinní podprostor A je nadrovina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$, kde předpokládáme $\mathbf{a} \neq \mathbf{0}$ (jinak by podprostor nebyl nadrovinou, tj. neměl by dimenzi $n - 1$). Jestliže $\|\mathbf{a}\| = 1$, vzdálenost bodu \mathbf{y} od nadroviny je $|\mathbf{a}^T \mathbf{y} - b|$. Obecně je tato vzdálenost rovna $|\mathbf{a}^T \mathbf{y} - b|/\|\mathbf{a}\|$. Vidíme zde geometrický význam vektoru \mathbf{a} a čísla b : \mathbf{a} je normálový vektor nadroviny a $|b|/\|\mathbf{a}\|$ je vzdálenost nadroviny od počátku.

4.7 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. Pro jaké vektory \mathbf{x}, \mathbf{y} platí trojúhelníková nerovnost $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ s rovností?
- 4.3. Najděte bázi ortogonálního doplňku prostoru $\text{span}\{(0, 1, 1), (1, 2, 3)\}$.
- 4.4. Jsou dány množiny $A = \text{span}\{(1, 0, 1, 0)\}$, $B = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_3 = 0\}$ a $C = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_3 = 0, x_2 = x_4\}$. Jsou to lineární podprostory? Pokud ano, najděte jejich dimenze. Rozhodněte, které z následujících výroků platí: $A \perp B$, $A \perp C$, $B \perp C$, $A = B^\perp$, $B = A^\perp$, $A = C^\perp$. Najděte libovolnou bázi podprostoru C^\perp .
- 4.5. Pro dva vektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ dokažte následující tvrzení, nakreslete obrázek a uvědomte si, jaké známé středoškolské poučky jste to vlastně dokázali.
- Jestliže $\|\mathbf{x}\| = \|\mathbf{y}\|$, pak $(\mathbf{x} + \mathbf{y}) \perp (\mathbf{x} - \mathbf{y})$.
 - Jestliže $\mathbf{x} \perp \mathbf{y}$, pak $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$.
 - Jestliže $\mathbf{x} \perp \mathbf{y}$, pak $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} + \mathbf{y}\|^2$.
 - Jestliže vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ jsou po dvojicích ortogonální (přesně: každá dvojice různých vektorů je ortogonální), pak $\|\mathbf{x}_1 + \dots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \dots + \|\mathbf{x}_k\|^2$.
- 4.6. Dle (4.16) a Věty 4.6 je každý podprostor komplementární se svým ortogonálním doplňkem. Dokažte, že to platí i obráceně: jestliže jsou dva podprostory ortogonální a komplementární, jsou ortogonálním doplňkem jeden druhého.
- 4.7. Najděte ortonormální bázi podprostoru $\text{span}\{(1, 1, 1, -1), (2, -1, -1, 1), (-1, 2, 2, 1)\}$ pomocí QR rozkladu (použijte Matlab).
- 4.8. Dokažte, že součin ortogonálních matic je ortogonální matice.
- 4.9. Pro jaké n je matice $\text{diag}(-\mathbf{1}_n)$ (tedy diagonální matice se samými mínus jedničkami na diagonále) rotační?
- 4.10. Za jakých podmínek na čísla a, b je matice $\begin{bmatrix} a+b & b-a \\ a-b & b+a \end{bmatrix}$ ortogonální?
- 4.11. Existuje isometrie $\mathbf{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ taková, že $\mathbf{f}(1, -1, 2) = (1, 2, -1, 1)$ a $\mathbf{f}(1, 1, 0) = (0, 1, -1, 0)$?
- 4.12. Isometrie (rotace+zrcadlení) v prostoru \mathbb{R}^n jsou reprezentovány ortogonálními maticemi. Pro $n > 3$ ovšem nemáme o rotacích a zrcadleních intuitivní představu.
- Počet nezávislých parametrů ('stupňů volnosti') ortogonální matice je rozdíl počtu prvků matice a počtu nezávislých (v našem případě různých) rovnic v podmínce $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Neformálně řečeno udává, kolika 'knoflíky' můžeme nezávisle 'kroutit' při rotaci v \mathbb{R}^n . Jaké je toto číslo pro $n = 2, 3, 4$? Najděte vzorec pro obecné n .
 - Je známo, že infinitezimálně malé rotace jsou v prvním řádu aproximovány maticí $\mathbf{I} + \mathbf{S}$ kde \mathbf{S} je antisymetrická matice. Ukažte, že to tak je. Spočítejte počet nezávislých parametrů antisymetrické matice $n \times n$ a ověřte, že je stejný jako v podúkolu a).
- 4.13. Máme vektory $\mathbf{x}_1 = (1, -1, 0, 2)$, $\mathbf{x}_2 = (1, 1, 1, 0)$, $\mathbf{x}_3 = (-1, -1, 2, 0)$.
- Ověřte, že vektory $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ jsou po dvojicích ortogonální.
 - Najděte libovolnou bázi podprostoru $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}^\perp$.
 - Najděte ortogonální projektory na podprostory $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ a $\text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}^\perp$.

- 4.14. Najděte dva ortogonální vektory \mathbf{x}, \mathbf{y} takové, že $\text{span}\{\mathbf{x}, \mathbf{y}\} = \text{span}\{(0, 1, 1), (1, 2, 3)\}$.
- 4.15. Spočítejte 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?
- 4.16. (★) 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:
- $X \subseteq Y \Rightarrow X^\perp \supseteq Y^\perp$.
 - $X + Y$ je generován sjednocením libovolné báze X a libovolné báze Y . Tedy je-li $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ báze X a $\{\mathbf{y}_1, \dots, \mathbf{y}_l\}$ báze Y , pak $X + Y = \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_l\}$.
 - $(X + Y)^\perp = X^\perp \cap Y^\perp$.
 - $(X \cap Y)^\perp = X^\perp + Y^\perp$.
- 4.17. Jak byste levně spočítali absolutní hodnotu determinantu matice z jejího QR rozkladu?
- 4.18. (★) *RQ rozklad* rozloží matici $\mathbf{A} = \mathbf{RQ}$, kde \mathbf{R} je horní trojúhelníková a \mathbf{Q} je ortogonální. Jak byste spočítali RQ rozklad z QR rozkladu?
- 4.19. Necht' $\mathbf{U}^T \mathbf{U} = \mathbf{I} = \mathbf{V}^T \mathbf{V}$. Dokažte, že $\text{rng } \mathbf{U} = \text{rng } \mathbf{V}$ (tedy sloupce \mathbf{U} a \mathbf{V} jsou ortonormální báze téhož podprostoru), právě když $\mathbf{V} = \mathbf{UC}$ pro nějakou ortogonální matici \mathbf{C} .
- 4.20. Z fyziky znáte vzorečky pro průmět $\frac{\mathbf{u}^T \mathbf{y}}{\|\mathbf{u}\|} \mathbf{u}$ a délku průmětu $\frac{\mathbf{u}^T \mathbf{y}}{\|\mathbf{u}\|}$ vektoru \mathbf{y} do směru daném vektorem \mathbf{u} (tedy do přímky $\text{span}\{\mathbf{u}\}$). Jak tyto vzorečky souvisejí s (4.23) a (4.27)?
- 4.21. Necht' $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Najděte co nejjednodušší vzorec pro
- vzdálenost bodu \mathbf{x} od podprostoru $\text{rng } \mathbf{U}$,
 - vzdálenost bodu \mathbf{x} od podprostoru $\text{null}(\mathbf{U}^T) = \{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{0}\}$,
 - vzdálenost počátku $\mathbf{0}$ od afinního podprostoru $\{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{b}\}$,
 - vzdálenost bodu \mathbf{x} od afinního podprostoru $\{\mathbf{x} \mid \mathbf{U}^T \mathbf{x} = \mathbf{b}\}$.
- 4.22. Necht' $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Dokažte, že
- Prvky matice \mathbf{U} splňují $|u_{ij}| \leq 1$.
 - Pro každý řádek \mathbf{u}^T matice \mathbf{U} platí $\|\mathbf{u}\| \leq 1$.
- 4.23. Necht' $\mathbf{P} \in \mathbb{R}^{n \times n}$ je ortogonální projektor. Dokažte, že
- Prvky matice \mathbf{P} splňují $|p_{ij}| \leq 1$.
 - Diagonální prvky matice \mathbf{P} jsou nezáporné.
 - $\|\mathbf{P}\|^2 = n$.
- 4.24. Afinní podprostor $A \subseteq \mathbb{R}^m$ lze reprezentovat jako $A = \mathbf{x} + \text{rng } \mathbf{U}$ kde $\mathbf{x} \in A$ a $\mathbf{U} \in \mathbb{R}^{m \times n}$. Bod \mathbf{x} ovšem není touto reprezentací určen jednoznačně: zvolíme-li libovolný jiný bod $\mathbf{y} \in A$, dle Věty 3.17 je $\mathbf{y} + \text{rng } \mathbf{U} = A$. Tuto nejednoznačnost můžeme odstranit přidáním podmínky, že norma $\|\mathbf{x}\|$ je minimální, tj. bod \mathbf{x} je průmětem počátku na podprostor A . Jsou-li dány $\mathbf{y} \in A$ a $\mathbf{U} \in \mathbb{R}^{m \times n}$, najděte co nejjednodušší vzorec pro takový bod \mathbf{x} .
- 4.25. Zobrazení $\mathbf{F}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ je dané vzorcem $\mathbf{F}(\mathbf{X}) = (\mathbf{I} - \mathbf{X})(\mathbf{I} + \mathbf{X})^{-1}$. Předpokládejte, že \mathbf{X} je taková, že $\mathbf{I} + \mathbf{X}$ je regulární. Dokažte, že:

- a) Pro každou \mathbf{X} je $(\mathbf{I} - \mathbf{X})(\mathbf{I} + \mathbf{X}) = (\mathbf{I} + \mathbf{X})(\mathbf{I} - \mathbf{X})$.
- b) Pro každou \mathbf{X} je $(\mathbf{I} - \mathbf{X})(\mathbf{I} + \mathbf{X})^{-1} = (\mathbf{I} + \mathbf{X})^{-1}(\mathbf{I} - \mathbf{X})$.
- c) Pro každou antisymetrickou \mathbf{X} je $\mathbf{F}(\mathbf{X})$ ortogonální.
- d) Pro každou ortogonální \mathbf{X} je $\mathbf{F}(\mathbf{X})$ antisymetrická.
- e) Zobrazení \mathbf{F} je inverzí sama sebe (tedy je to *involve*, viz §1.1.2), tj. $\mathbf{F}(\mathbf{F}(\mathbf{X})) = \mathbf{X}$ pro každou \mathbf{X} .

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

Nápověda a řešení

- 4.1. (a) $\|\mathbf{x}\| = \sqrt{14}$, (b) $\|\mathbf{x} - \mathbf{y}\| = 2\sqrt{3}$, (c) ≈ 1.1832 radiánů
- 4.3. Např. $(1, 1, -1)$
- 4.5.a) Dokazujeme, že úhlopříčky kosočtverce jsou na sebe kolmé.
- 4.5.b) $\|\mathbf{x} - \mathbf{y}\|^2 = (\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y}) = \mathbf{x}^T\mathbf{x} - \mathbf{y}^T\mathbf{x} - \mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{y} = \|\mathbf{x}\|^2 - 2\mathbf{x}^T\mathbf{y} + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, protože $\mathbf{x}^T\mathbf{y} = 0$. Dokázali jsme Pythagorovu větu.
- 4.5.d) Dokazované tvrzení je zobecnění Pythagorovy věty.
- 4.7. Báze je $\{(1, 1, 1, -1)/2, (3, -1, -1, 1)/\sqrt{12}, (0, 1, 1, 2)/\sqrt{6}\}$
- 4.8. Necht' $\mathbf{A}^T\mathbf{A} = \mathbf{I} = \mathbf{B}^T\mathbf{B}$ a $\mathbf{C} = \mathbf{AB}$. Pak $\mathbf{C}^T\mathbf{C} = \mathbf{B}^T\mathbf{A}^T\mathbf{AB} = \mathbf{B}^T\mathbf{B} = \mathbf{I}$.
- 4.9. Musí být $\det \text{diag}(-\mathbf{1}_n) = (-1)^n > 0$, tedy pro sudá n .
- 4.11. Ne, protože isometrie zachovává eukleidovskou normu, ale $\|(1, -1, 2)\| \neq \|(1, 2, -1, 1)\|$.
- 4.12.a) Podmínka $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ je soustava rovnic $\mathbf{u}_i^T\mathbf{u}_j = \delta_{ij}$ pro $i, j = 1, \dots, n$, kde $\mathbf{u}_1, \dots, \mathbf{u}_n$ jsou sloupce matice \mathbf{U} . Tato soustava obsahuje jen $\binom{n}{2} + n$ různých rovnic (napište si ji např. pro $n = 4$). Tedy počet stupňů volnosti je $n^2 - \binom{n}{2} - n = \binom{n}{2} = n(n-1)/2$.
- 4.12.b) Matice $\mathbf{I} + t\mathbf{S}$ pro velmi malé t musí být v prvním řádu blízká rotační.
- 4.14. 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.)
- 4.15. Není matice náhodou ortogonální? Ort. projektor to není, protože nesplňuje $\mathbf{P}^2 = \mathbf{P}$.
- 4.16.c) Plyne z (b).
- 4.16.d) Plyne z (c) s použitím $(X^\perp)^\perp = X$.
- 4.19. Dokážeme jen jednu implikaci. Víme, že $\mathbf{U} \in \mathbb{R}^{m \times n}$ a $\mathbf{C} \in \mathbb{R}^{n \times n}$, kde $m \geq n$, a $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{C}^T\mathbf{C} = \mathbf{I} = \mathbf{CC}^T$. Sloupce matic \mathbf{U} a \mathbf{V} jsou báze stejného podprostoru, neboť dle Tvzení 3.12 je $\text{rng } \mathbf{U} = \text{rng } \mathbf{V}$ (\mathbf{C} je matice přechodu k jiné bázi). Sloupce matice \mathbf{V} jsou ortonormální, neboť $\mathbf{V}^T\mathbf{V} = \mathbf{C}^T\mathbf{U}^T\mathbf{UC} = \mathbf{C}^T\mathbf{C} = \mathbf{I}$.
- 4.20. Matice \mathbf{U} v (4.23) a (4.27) je normalizovaný vektor \mathbf{u} , tedy vektor $\mathbf{U} = \mathbf{u}/\|\mathbf{u}\|$.
- 4.21.a) $\|(\mathbf{I} - \mathbf{UU}^T)\mathbf{x}\|$, víc to už zjednodušit nejde.
- 4.21.b) $\|\mathbf{U}^T\mathbf{x}\|$, viz §4.6.2.
- 4.21.c) $\|\mathbf{b}\|$, viz §4.6.2.
- 4.21.d) $\|\mathbf{U}^T\mathbf{x} - \mathbf{b}\|$, viz §4.6.2.
- 4.22.b) Doplníme matici \mathbf{U} na ortogonální matici $\mathbf{W} = [\mathbf{U} \quad \mathbf{V}]$. Řádek této matice je $\mathbf{w}^T = [\mathbf{u}^T \quad \mathbf{v}^T]$. Z ortogonality \mathbf{W} je však $\mathbf{w}^T\mathbf{w} = \mathbf{u}^T\mathbf{u} + \mathbf{v}^T\mathbf{v} = 1$. Z toho $\mathbf{u}^T\mathbf{u} \leq 1$.
- 4.23.a) Z $\mathbf{P} = \mathbf{UU}^T$ 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 $|\mathbf{u}_i^T\mathbf{u}_j| \leq 1$.
- 4.23.b) Máme $p_{ii} = \mathbf{u}_i^T\mathbf{u}_i \geq 0$

4.23.c) Použijte vzorec $\|\mathbf{P}\|^2 = \text{tr}(\mathbf{P}^T\mathbf{P})$, dosad'te za \mathbf{P} a upravujte.

4.24. $\mathbf{x} = (\mathbf{I} - \mathbf{U}\mathbf{U}^T)\mathbf{y}$

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). Množina řešení soustavy je afinní podprostor \mathbb{R}^n (dle Věty 3.18).

V této kapitole se zaměříme pouze na nehomogenní soustavy (tj. $\mathbf{b} \neq \mathbf{0}$). 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

Když soustava (5.1) nemá řešení, často je užitečné nalézt takové \mathbf{x} , aby rovnost (5.1) platila aspoň přibližně (což můžeme zapsat jako $\mathbf{Ax} \approx \mathbf{b}$). Přesněji, 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 přeurlčená. Její přibližné řešení ve smyslu nejmenších čtverců znamená najít taková čísla x, y , která minimalizují číslo $(x + 2y - 6)^2 + (-x + y - 3)^2 + (x + y - 4)^2$. ♦

Věta 5.1. Vektor \mathbf{x} je optimální řešení úlohy (5.2), právě když splňuje

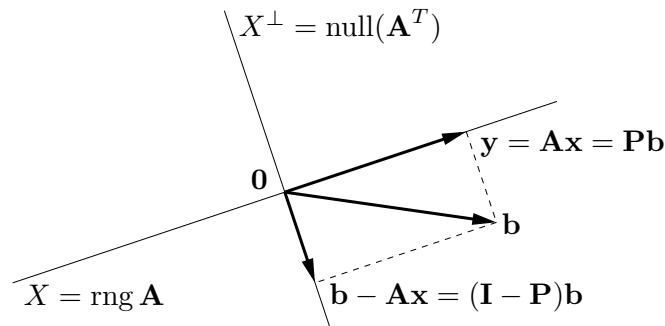
$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}. \quad (5.3)$$

Důkaz. Substitucí $\mathbf{y} = \mathbf{A} \mathbf{x}$ lze úlohu (5.2) přepsat na (viz definice (3.16) prostoru obrazů)

$$\min_{\mathbf{y} \in \text{rng } \mathbf{A}} \|\mathbf{y} - \mathbf{b}\|^2. \quad (5.4)$$

Dle Věty 4.11 je vzdálenost $\|\mathbf{y} - \mathbf{b}\|$ minimální právě když $(\mathbf{b} - \mathbf{y}) \perp \text{rng } \mathbf{A}$, neboli \mathbf{y} je ortogonální projekce bodu \mathbf{b} na $\text{rng } \mathbf{A}$. To znamená (viz (4.15)), že vektor $\mathbf{b} - \mathbf{A} \mathbf{x}$ je kolmý na každý sloupec matice \mathbf{A} . Tato podmínka se zapíše jako $\mathbf{A}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) = \mathbf{0}$, což je (5.3). ■

Vše je vidět na obrázku, který je vlastně překreslený obrázek z §4.6 v jiném značení (např. \mathbf{x} znamená něco jiného v každém obrázku; porovnejte obrázky a promyslete!):



Soustava (5.3) se nazývá soustava **normálních rovnic** (protože normála = kolmice). Je to soustava n rovnic o n neznámých. Abychom mohli zkoumat její řešitelnost, potřebujeme dokázat důležitou vlastnost matic tvaru² $\mathbf{A}^T \mathbf{A}$ a $\mathbf{A} \mathbf{A}^T$.

Věta 5.2. Pro každou matici \mathbf{A} platí

$$\text{rng}(\mathbf{A}^T \mathbf{A}) = \text{rng}(\mathbf{A}^T), \quad (5.5a)$$

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

Důkaz. Dokažme nejprve rovnost (5.5b), tj. $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{A} \mathbf{x} = \mathbf{0}$. Implikace \Leftarrow je snadná, vynásobením $\mathbf{A} \mathbf{x} = \mathbf{0}$ zleva maticí \mathbf{A}^T . Implikace \Rightarrow se dokáže takto:

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

neboť pro libovolný vektor \mathbf{y} platí $\mathbf{y}^T \mathbf{y} = 0 \Rightarrow \mathbf{y} = \mathbf{0}$ (proč?).

Dokažme (5.5a). Z definice (3.16) je jasné (viz Tvrzení 3.12), že $\text{rng}(\mathbf{A}^T \mathbf{A}) \subseteq \text{rng}(\mathbf{A}^T)$. Nyní použijeme Větu 3.13 jednou na matici \mathbf{A} a jednou na matici $\mathbf{A}^T \mathbf{A}$:

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

Díky (5.5b) z toho máme $\dim \text{rng}(\mathbf{A}^T \mathbf{A}) = \dim \text{rng } \mathbf{A} = \dim \text{rng}(\mathbf{A}^T)$, kde druhá rovnost plyne z Věty 3.9. Z Věty 3.3 tedy máme (5.5a). ■

²Maticе tvaru $\mathbf{A}^T \mathbf{A}$ či $\mathbf{A} \mathbf{A}^T$ se často 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 (2.21)). Matici $\mathbf{A} \mathbf{A}^T \in \mathbb{R}^{m \times m}$ lze zase vidět (až na skalární násobek) jako empirickou *kovarianční matici* n pozorování $\mathbf{a}_1, \dots, \mathbf{a}_n$ m -tice náhodných proměnných.

Důsledek 5.3. *Soustava (5.3) má (aspoň jedno) řešení pro každé \mathbf{A} , \mathbf{b} .*

Důkaz. Soustava (5.3) má řešení právě tehdy, když $\mathbf{A}^T \mathbf{b} \in \text{rng}(\mathbf{A}^T \mathbf{A}) = \text{rng}(\mathbf{A}^T)$, kde rovnost je (5.5a). Ale z definice prostoru obrazů je $\mathbf{A}^T \mathbf{b} \in \text{rng}(\mathbf{A}^T)$ pro libovolná \mathbf{A} , \mathbf{b} . ■

Zkombinujeme-li (5.5a) a Větu 3.9, máme

$$\text{rank}(\mathbf{A}\mathbf{A}^T) = \text{rank } \mathbf{A} = \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A}^T \mathbf{A}). \quad (5.6)$$

Dle (5.6) 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

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

Matice (5.7) 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 (5.7) nelze použít (zkrátka proto, že matice $\mathbf{A}^T \mathbf{A}$ nemá inverzi). Potom 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 jsou báze podprostoru X), z (5.7) máme $\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$, kde

$$\mathbf{P} = \mathbf{A}\mathbf{A}^+ = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T. \quad (5.8)$$

Toto je tedy projektor na podprostor X s bází (ne nutně ortonormální) tvořenou sloupci matice \mathbf{A} . Zdůrazněme, že projektor (5.8) vyjde stejný pro libovolnou bázi podprostoru X , což plyne z Věty 3.15 (a elementárně to máte dokázat ve Cvičení 5.11). Pokud je báze ortonormální, tj. $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, pak se (5.8) redukuje na (4.21).

Projekce na X^\perp má přirozenou roli v úloze (5.2): je-li \mathbf{x} optimální řešení úlohy, pak její optimální hodnota 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 (5.7) 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 $\mathbf{A}^T \mathbf{A}$ k zaokrouhlení a výsledná matice je singulární. Tedy soustava $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}$ nemá řešení. \blacklozenge

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

$$\mathbf{R} \mathbf{x} = \mathbf{Q}^T \mathbf{b}. \quad (5.9)$$

Zdůraněme, že pokud \mathbf{A} není čtvercová, pak soustava (5.9) není ekvivalentní původní soustavě $\mathbf{A} \mathbf{x} = \mathbf{b}$.

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 přeürčená, 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. Modelujeme závislost proměnné $y \in \mathbb{R}$ na proměnné $x \in X$ (kde X je libovolná množina) 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

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^n} \sum_{i=1}^m (y_i - f(x_i, \boldsymbol{\theta}))^2. \quad (5.10)$$

Často je regresní funkce taková, že pro každé x je lineární funkcí parametrů $\boldsymbol{\theta}$. V tom případě mluvíme o **lineární regresi**. Taková funkce jde vždy napsat jako lineární kombinace

$$f(x, \boldsymbol{\theta}) = \theta_1 \varphi_1(x) + \dots + \theta_n \varphi_n(x) = \boldsymbol{\varphi}(x)^T \boldsymbol{\theta} \quad (5.11)$$

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

$$\mathbf{A} = \begin{bmatrix} \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_m) & \dots & \varphi_n(x_m) \end{bmatrix}$$

Tedy převedli jsme úlohu (5.10) na tvar (5.2).

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

Příklad 5.3. Nejjednodušší případ je pro $n = 1$ a konstantní funkci $\varphi_i(x) = 1$. Funkce (5.11) je tedy $f(x, \theta) = \theta$. Úloha (5.10) 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 . ♦

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

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

je polynom stupně $n - 1$ proměnné x . Matice

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

je známá jako *Vandermondova matice*.

Tento případ jde snadno zobecnit na polynomy více proměnných: máme $X = \mathbb{R}^d$ a bázové funkce jsou monomy proměnných x_1, \dots, x_d (viz §6) až do nějakého stupně. ♦

5.1.4 Statistické odůvodnění kritéria nejmenších čtverců

Možná se ptáte, proč se má nalezení přibližného řešení přeurčené 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

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{r}, \quad (5.12)$$

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

$$p(\mathbf{r}) = p(r_1, \dots, r_m) = \prod_{i=1}^m p(r_i) = \prod_{i=1}^m c e^{-r_i^2/(2\sigma^2)}. \quad (5.13)$$

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 $\sum_i r_i^2$.

⁴Nedejte se zmást tím, že polynom není lineární funkce a přesto jde o lineární regresi. Důležité je, že regresní funkce (5.11) je lineární v parametrech $\boldsymbol{\theta}$.

5.1.5 Vícekriteriální nejmenší čtverce, regularizace

V některých úlohách se hodí minimalizovat více kritérií tvaru $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|$ ‘současně’. K tomu se dá přistoupit tak, že minimalizujeme (nezáporně) vážený součet kritérií⁵, tedy funkci⁶

$$\mu_1\|\mathbf{A}_1\mathbf{x} - \mathbf{b}_1\|^2 + \cdots + \mu_k\|\mathbf{A}_k\mathbf{x} - \mathbf{b}_k\|^2 \quad (5.14)$$

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 (5.14) je roven (viz Cvičení 5.16)

$$\left\| \begin{bmatrix} \sqrt{\mu_1}(\mathbf{A}_1\mathbf{x} - \mathbf{b}_1) \\ \vdots \\ \sqrt{\mu_k}(\mathbf{A}_k\mathbf{x} - \mathbf{b}_k) \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \sqrt{\mu_1}\mathbf{A}_1 \\ \vdots \\ \sqrt{\mu_k}\mathbf{A}_k \end{bmatrix} \mathbf{x} - \begin{bmatrix} \sqrt{\mu_1}\mathbf{b}_1 \\ \vdots \\ \sqrt{\mu_k}\mathbf{b}_k \end{bmatrix} \right\|^2 = \|\mathbf{A}'\mathbf{x} - \mathbf{b}'\|^2, \quad (5.15)$$

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

$$\mathbf{x} = (\mathbf{A}'^T \mathbf{A}')^{-1} \mathbf{A}'^T \mathbf{b}' = (\mu_1 \mathbf{A}_1^T \mathbf{A}_1 + \cdots + \mu_k \mathbf{A}_k^T \mathbf{A}_k)^{-1} (\mu_1 \mathbf{A}_1^T \mathbf{b}_1 + \cdots + \mu_k \mathbf{A}_k^T \mathbf{b}_k). \quad (5.16)$$

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

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \mu\|\mathbf{x}\|^2). \quad (5.17)$$

pro zvolenou váhu $\mu > 0$. Přidání členu $\mu\|\mathbf{x}\|^2$ se říká (Tichonovova) **regularizace** úlohy (5.2). Dosazením do vzorečku (5.16) ukážeme (proved'te!), že optimální řešení je rovno $\mathbf{x} = \mathbf{A}_\mu^+ \mathbf{b}$ kde

$$\mathbf{A}_\mu^+ = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T \quad (5.18)$$

je ‘regularizovaná pseudoinverze’⁷ 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 5.17), tedy \mathbf{A}_μ^+ je vždy definována.

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

$$\min\{\|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}. \quad (5.19)$$

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 11.18.

⁵Minimalizací více kritérií současně se zabývá obor *vícekriteriální optimalizace* (*multiobjective optimization*).

⁶Matematicky 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 (5.14) by pak bylo kratší.

⁷Symbol \mathbf{A}_μ^+ zde neoznačuje pseudoinverzi nějaké matice \mathbf{A}_μ , taková matice \mathbf{A}_μ totiž neexistuje.

Příklad 5.5. Soustava dvou rovnic o třech neznámých

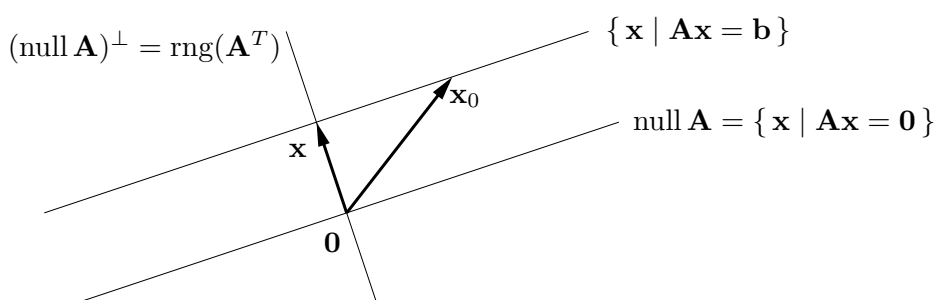
$$\begin{aligned} x + 2y + z &= 1 \\ -x + y + 2z &= 2 \end{aligned}$$

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

Množinu řešení soustavy (5.1) lze psát (viz důkaz Věty 3.18) jako

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \} = \text{null } \mathbf{A} + \mathbf{x}_0, \quad (5.20)$$

kde \mathbf{x}_0 je libovolné (partikulární) řešení soustavy, tedy $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$. Množina (5.20) 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 $\mathbf{x} \perp \text{null } \mathbf{A}$, tj. $\mathbf{x} \in (\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T)$ (viz (4.19b)). Neboli musí existovat vektor $\mathbf{y} \in \mathbb{R}^m$ tak, že $\mathbf{x} = \mathbf{A}^T \mathbf{y}$. Pro vyřešení úlohy (5.19) tedy musíme vyřešit soustavu rovnic

$$\mathbf{A}^T \mathbf{y} = \mathbf{x}, \quad (5.21a)$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \quad (5.21b)$$

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

Vyřešíme tuto soustavu za předpokladu, že matice \mathbf{A} má lineárně nezávislé řádky. Dosazením \mathbf{x} do druhé rovnice obdržíme $\mathbf{A}\mathbf{A}^T \mathbf{y} = \mathbf{b}$. Tato soustava má řešení, jestliže má soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ řešení, protože pak $\mathbf{b} \in \text{rng } \mathbf{A} = \text{rng}(\mathbf{A}\mathbf{A}^T)$. Protože \mathbf{A} má lineárně nezávislé řádky, dle (5.6) je matice $\mathbf{A}\mathbf{A}^T$ regulární a tedy $\mathbf{y} = (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b}$. Dosazením do první rovnice dostaneme $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, kde

$$\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \quad (5.22)$$

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 jinak. 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 (5.8), zde ovšem promítáme na $\text{rng}(\mathbf{A}^T)$ a tedy musíme vzorec použít s \mathbf{A}^T místo s \mathbf{A} . Tedy

$$\mathbf{x} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A}\mathbf{x}_0 = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b} = \mathbf{A}^+ \mathbf{b}. \quad (5.23)$$

Tvrzení 5.4. Je-li úloha (5.19) přípustná (tj. soustava $\mathbf{Ax} = \mathbf{b}$ má aspoň jedno řešení), pak tato úloha má právě jedno optimální řešení.

Důkaz. Výše jsme ukázali, že věta platí jestliže \mathbf{A} má lineárně nezávislé řádky. Ale pokud \mathbf{A} má lineárně závislé řádky, vynecháním vhodných rovnic ze soustavy $\mathbf{Ax} = \mathbf{b}$ lze učinit řádky \mathbf{A} lineárně nezávislé při zachování množiny řešení soustavy. ■

Vzorce (5.7) a (5.22) dohromady definují pseudoinverzi libovolné matice (čtvercové, úzké nebo široké) s plnou hodností (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

$$\min\{\|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}\}. \quad (5.24)$$

Podle Tvrzení 5.3 a 5.4 má tato úloha právě jedno optimální řešení pro libovolná \mathbf{A}, \mathbf{b} . Toto řešení, \mathbf{x}^* , má následující vlastnosti:

- Má-li soustava $\mathbf{Ax} = \mathbf{b}$ jediné řešení, \mathbf{x}^* je toto řešení.
- Nemá-li soustava $\mathbf{Ax} = \mathbf{b}$ řešení, \mathbf{x}^* je její přibližné řešení ve smyslu nejmenších čtverců, tj. optimální řešení úlohy (5.2). Pokud ovšem úloha (5.2) má více než jedno (tedy nekonečně mnoho) optimálních řešení, \mathbf{x}^* je optimální řešení úlohy (5.2) které má nejmenší normu.
- Má-li soustava $\mathbf{Ax} = \mathbf{b}$ nekonečně mnoho řešení, \mathbf{x}^* je řešení této soustavy s nejmenší normou, tj. optimální řešení problému (5.19).

Podle §5.2 je \mathbf{x} optimální řešení úlohy (5.24) právě tehdy, když existuje \mathbf{y} takové, že

$$\mathbf{A}^T \mathbf{Ay} = \mathbf{x}, \quad (5.25a)$$

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}. \quad (5.25b)$$

Tuto soustavu obecně nejde řešit pomocí inverze, protože \mathbf{A} nemusí mít plnou hodnost. Z tvaru soustavy (5.25) ale plyne, že $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$ pro nějakou matici \mathbf{A}^+ která nezávisí na \mathbf{b} . Tutu matici nazýváme **pseudoinverze** (přesněji *Moore-Penroseova pseudoinverze*) matice \mathbf{A} . Když \mathbf{A} má lineárně nezávislé sloupce, \mathbf{A}^+ je (5.7). Když \mathbf{A} má lineárně nezávislé řádky, \mathbf{A}^+ je (5.22). Když ovšem \mathbf{A} nemá plnou hodnost, \mathbf{A}^+ je nutno počítat jinak. Elegantně se to udělá pomocí singulárního rozkladu (SVD), což ukážeme už zde (vrat' se k tomu po přečtení §7.4):

Věta 5.5. Necht' $\mathbf{A} = \mathbf{USV}^T$ je kompaktní SVD matice \mathbf{A} , tj. \mathbf{U}, \mathbf{V} mají ortonormální sloupce a \mathbf{S} je diagonální regulární. Pak

$$\mathbf{A}^+ = \mathbf{VS}^{-1}\mathbf{U}^T. \quad (5.26)$$

Důkaz. Stačí dokázat, že $\mathbf{x} = \mathbf{A}^+ \mathbf{b} = \mathbf{VS}^{-1}\mathbf{U}^T \mathbf{b}$ a $\mathbf{y} = \mathbf{VS}^{-2}\mathbf{V}^T \mathbf{x}$ splňují (5.25):

$$\begin{aligned} \mathbf{A}^T \mathbf{Ax} &= \mathbf{VS}^T \mathbf{U}^T \mathbf{USV}^T = \mathbf{VS}^2 \mathbf{V}^T, \\ \mathbf{A}^T \mathbf{Ay} &= \mathbf{VS}^2 \mathbf{V}^T \mathbf{y} = \mathbf{VS}^2 \mathbf{V}^T \mathbf{VS}^{-2} \mathbf{V}^T \mathbf{x} = \mathbf{x}, \\ \mathbf{A}^T \mathbf{Ax} &= \mathbf{VS}^2 \mathbf{V}^T \mathbf{x} = \mathbf{VS}^2 \mathbf{V}^T \mathbf{VS}^{-1} \mathbf{U}^T \mathbf{b} = \mathbf{VSU}^T \mathbf{b} = \mathbf{A}^T \mathbf{b}. \end{aligned} \quad \blacksquare$$

Zmíníme ještě jednu zajímavou úvahu, jak dospět k pseudoinverzi obecné matice. Všimněte si, že úloha (5.17) je ‘něco mezi’ úlohami (5.2) a (5.19). Neformálně, (5.19) si lze představit jako minimalizaci $\|\mathbf{x}\|^2 + \mu\|\mathbf{Ax} - \mathbf{b}\|^2$ pro velmi velké μ . To je ale totéž jako (5.17) pro velmi malé (kladné) μ . Lze dokázat (což uděláte ve Cvičení 7.20, pro další vhléd viz Cvičení 5.18), že pseudoinverze obecné matice je rovna

$$\mathbf{A}^+ = \lim_{\mu \rightarrow 0^+} \mathbf{A}_\mu^+. \quad (5.27)$$

(Moore-Penroseova) pseudoinverze obecné matice \mathbf{A} se v lineární algebře obvykle definuje jako matice jednoznačně určená podmínkami

$$(\mathbf{AA}^+)^T = \mathbf{AA}^+, \quad (\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}, \quad \mathbf{AA}^+\mathbf{A} = \mathbf{A}, \quad \mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+. \quad (5.28)$$

To nebudeme dokazovat, ale aspoň ověřte, že matice (5.7), (5.22) a (5.26) tyto podmínky splňují (více viz Cvičení 5.9). Pro praktické řešení lineárních soustav se pseudoinverze obecné matice moc nepoužívá, protože \mathbf{A}^+ je nespojitá funkce matice \mathbf{A} .

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.

- Příklad 1.12.
- Hledá se vzdálenost bodu $\mathbf{y} \in \mathbb{R}^n$ od přímky $\{\mathbf{a} + t\mathbf{s} \mid t \in \mathbb{R}\}$ kde $\mathbf{a}, \mathbf{s} \in \mathbb{R}^n$.
- Příklad 1.9.
- Máme množinu m přímek v \mathbb{R}^n , kde i -tá přímka je množina $\{\mathbf{a}_i + t\mathbf{s}_i \mid t \in \mathbb{R}\}$ pro dané $\mathbf{a}_i, \mathbf{s}_i \in \mathbb{R}^n$. Hledá se bod $\mathbf{y} \in \mathbb{R}^n$, jehož součet čtverců vzdáleností k přímkám je minimální.
- Máme m nadrovin v prostoru \mathbb{R}^n , kde i -tá nad rovina má rovnici $\mathbf{a}_i^T \mathbf{x} = b_i$ pro dané $\mathbf{a}_i \in \mathbb{R}^n$ a $b_i \in \mathbb{R}$. Hledá se bod $\mathbf{y} \in \mathbb{R}^n$, který minimalizuje součet čtverců vzdáleností od jednotlivých nadrovin.

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 interpretujte.
2. Jsou sloupce \mathbf{P} lineárně nezávislé?

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

- g) Závislost výkonu P kotle na průtoku G plynu a průřezu S díry na přívod vzduchu je modelována funkcí $\hat{P}(G, S) = G(a_1 + a_2S + a_310^{G+S} + a_410^{-S})$. 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 zaved'te diagonální matici $\mathbf{W} = \text{diag}(w_1, \dots, w_m)$. Napište normální rovnici a pseudoinverzi pro tento případ.

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

5.6. Necht' $X = \text{span}\left\{ \left(-\frac{3}{5}, 0, \frac{4}{5}, 0\right), (0, 0, 0, 1), \left(\frac{4}{5}, 0, \frac{3}{5}, 0\right) \right\}$. 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 (5.8) a podprostor $X = \text{rng } \mathbf{A}$? Jaký je geometrický význam této situace?

5.9. Dokažte následující vlastnosti pseudoinverze ze vztahů (5.7) a (5.22) 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) $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$, $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$
- f) $\mathbf{A}^T = \mathbf{A}^T\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+\mathbf{A}\mathbf{A}^T$

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

- 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 (5.8). Báze podprostoru X , na který promítáme, jsou sloupce matice \mathbf{A} . Projektor \mathbf{P} se nesmí změnit, vezmeme-li jinou bázi podprostoru. Napovíme (rozmyslete důkaz!), že všechny možné báze podprostoru X jsou dány sloupci matice $\mathbf{A} \mathbf{C}$ pro všechny možné regulární matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ (tedy \mathbf{C} je matice přechodu k jiné bázi).
- 5.12. Jaké bude řešení normálních rovnic (5.3) v případě, že \mathbf{A} má lineárně nezávislé sloupce a $\mathbf{b} \perp \text{rng } \mathbf{A}$? Vyřešte geometrickou úvahou (vzpomeňte si na ortogonální projekci a koukejte na obrázek v §5.1!) 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. 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 $(\text{rng } \mathbf{A})^\perp = \text{null } \mathbf{A}$.
- 5.16. Pro vektory $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ ukažte, že $\left\| \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \right\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2$.
- 5.17. Dokažte, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ je regulární pro každou matici \mathbf{A} a každé $\mu > 0$.
- 5.18. Uvažujme ‘regularizovanou pseudoinverzi’ (5.18). 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.
- 5.19. (*) 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 Tvzení 4.10.

Nápověda a řešení

5.1.a) Neplatí. Příklad: $m = 1, n = 2, \mathbf{A} = [0 \ 0], \mathbf{b} = 1$.

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 $\text{rng } \mathbf{A} = \mathbb{R}^m$, tedy soustava má řešení. Navíc má \mathbf{A} netriviální nulový prostor, tedy má nekonečně mnoho řešení.

5.2. $(x_1, x_2, x_3) = (2, 1, 0)/3$

5.3.a) Zbývá napsat funkci (1.20) ve tvaru $f(\mathbf{u}) = \|\mathbf{P} \mathbf{u} - \mathbf{q}\|^2$. To je snadné (srov. s (5.15)): máme

$$\sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|^2 = \left\| \begin{bmatrix} \mathbf{x} - \mathbf{a}_1 \\ \vdots \\ \mathbf{x} - \mathbf{a}_m \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \mathbf{I}_n \\ \vdots \\ \mathbf{I}_n \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \right\|^2, \text{ tedy } \mathbf{u} = \mathbf{x}, \mathbf{P} = \begin{bmatrix} \mathbf{I}_n \\ \vdots \\ \mathbf{I}_n \end{bmatrix}, \mathbf{q} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$$

5.3.c) Zbývá napsat funkci (1.14) ve tvaru $f(\mathbf{u}) = \|\mathbf{P} \mathbf{u} - \mathbf{q}\|^2$. To je opět snadné: $\mathbf{u} = (\alpha_1, \alpha_2)$, $\mathbf{P} = [\mathbf{b}_1 - \mathbf{a}_1 \ \mathbf{a}_2 - \mathbf{b}_2], \mathbf{q} = \mathbf{a}_1 - \mathbf{a}_2$.

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 (5.7) 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 (5.22) 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 (5.5a) je $\text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{B}^T)$. Tedy každý řádek \mathbf{A} je lineární kombinací řádků \mathbf{B} . Dle Věty 3.8 to jde napsat jako $\mathbf{A} = \mathbf{C}\mathbf{B}$ pro nějaké \mathbf{C} .

5.14.a) $|b|/\|\mathbf{a}\|$

5.14.b) Čtverec vzdálenosti je roven optimální hodnotě úlohy (5.19), tedy $(\mathbf{A}^+ \mathbf{b})^T (\mathbf{A}^+ \mathbf{b}) = \mathbf{b}^T (\mathbf{A} \mathbf{A}^T)^{-T} \mathbf{A} \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{b} = \mathbf{b}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{b}$.

5.14.c) $|\mathbf{a}^T \mathbf{x} - b|/\|\mathbf{a}\|$

5.15. Dle (4.19a) a (5.5b) 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. 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 (5.6) má $\mathbf{B}^T \mathbf{B}$ plnou hodnotu, tedy je regulární.

5.18. 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.

5.19. $\mathbf{I} = [\mathbf{A} \ \mathbf{B}] [\mathbf{A} \ \mathbf{B}]^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} = [\mathbf{A} \ \mathbf{B}] \left(\begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} [\mathbf{A} \ \mathbf{B}] \right)^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}$
 $= [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^T \mathbf{B} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} = [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} (\mathbf{A}^T \mathbf{A})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{B}^T \mathbf{B})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}$
 $= \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T + \mathbf{B} (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$.

Význam: $[\mathbf{A} \ \mathbf{B}]$ 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** (angl. *monomial*) 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** (angl. *polynomial*) n proměnných je lineární kombinace monomů, přičemž **stupeň polynomu** je stupeň jeho monomu (s nenulovým koeficientem) 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 (3.37) je jen jiný název pro polynom prvního stupně a lineární funkce (3.11) (také zvaná lineární forma) je jiný název pro homogenní polynom prvního stupně¹. Polynom druhého stupně se nazývá *kvadratická funkce* a homogenní polynom druhého stupně *kvadratická forma*². Cílem této kapitoly je porozumět extrémům kvadratických forem a funkcí.

6.1 Vlastní čísla a vektory

Nechť pro čtvercovou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$, *nenulový* vektor $\mathbf{v} \in \mathbb{C}^n$ a skalár $\lambda \in \mathbb{C}$ platí³

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

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, i když \mathbf{A} má všechny prvky reálné. Množině všech vlastních čísel matice se říká také její *spektrum*.

Rovnost (6.2) je soustava n *nelineárních* rovnic s $n + 1$ neznámými λ, \mathbf{v} . Z této soustavy lze eliminovat \mathbf{v} známým obratem, čímž dostaneme jedinou rovnici pro λ . Přepíšeme (6.2) jako

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

¹Tohle platí jen pro funkce na lineárním prostoru \mathbb{R}^n . I když polynomy lze definovat (i když složitějším způsobem) na abstraktním (tedy definovaným axiomy) lineárním prostoru, obvykle se jim tak neříká.

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

³Zopakujme, že \mathbb{C} je množina komplexních čísel, tj. čísel ve tvaru $a + bi$ kde $a, b \in \mathbb{R}$ a $i = \sqrt{-1}$.

což pro pevné λ je soustava homogenních lineárních rovnic pro \mathbf{v} . Protože vlastní vektory nesmí být nulové, λ splňuje (6.2) právě tehdy, když soustava (6.3) má netriviální (tj. nenulové) řešení. To nastane právě tehdy, když její matice $\mathbf{A} - \lambda\mathbf{I}$ je singulární, neboli

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

Dle definice determinantu (2.5) je funkce $p_{\mathbf{A}}: \mathbb{R} \rightarrow \mathbb{R}$ polynom stupně n . Nazývá se **charakteristický polynom** matice \mathbf{A} . Podle tzv. *základní věty algebry* má každý polynom stupně n právě n komplexních kořenů, počítáme-li každý kořen tolikrát, kolik je jeho násobnost. Proto:

Věta 6.1. Každá matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ má nejméně jedno a nejvýše n různých vlastních čísel.

Označíme-li různé kořeny charakteristického polynomu jako $\lambda_1, \dots, \lambda_m$ (kde $m \leq n$), je tedy

$$p_{\mathbf{A}}(\lambda) = \prod_{i=1}^m (\lambda_i - \lambda)^{n_i} \quad \text{kde} \quad \sum_{i=1}^m n_i = n. \quad (6.5)$$

Zde m je počet různých kořenů charakteristického polynomu (tedy různých vlastních čísel matice) a n_i je násobnost kořene λ_i . Číslo n_i se nazývá **algebraická násobnost** vlastního čísla λ_i .

Je-li λ vlastní číslo, pak jemu příslušné vlastní vektory tvoří množinu řešení homogenní lineární soustavy (6.3), kromě počátku $\mathbf{0}$ (protože vlastní vektory nesmí být nulové). Tato množina tvoří podprostor $\text{null}(\mathbf{A} - \lambda\mathbf{I})$, kterému říkáme **vlastní podprostor** příslušný vlastnímu číslu λ . Dimenze tohoto podprostoru se nazývá **geometrická násobnost** vlastního čísla λ .

Příklad 6.1. Vlastní čísla matice $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$ jsou řešeními rovnice

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

Tato kvadratická rovnice má dva kořeny, $\lambda_1 = -1$ a $\lambda_2 = -2$. To jsou tedy dvě vlastní čísla matice, každé s algebraickou násobností 1.

Vlastní vektor příslušný vlastnímu číslu λ_1 najdeme řešením homogenní lineární soustavy

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

Jeden takový vektor je $\mathbf{v}_1 = (-1, 1)$. Vlastní podprostor příslušný λ_1 je přímka $\text{span}\{\mathbf{v}_1\} = \{\alpha(-1, 1) \mid \alpha \in \mathbb{R}\}$. Má dimenzi 1, což je tedy geometrická násobnost vlastního čísla λ_1 . ♦

Příklad 6.2. Matice $\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ má jediné vlastní číslo $\lambda = 0$ s algebraickou násobností 2. Jemu příslušný vlastní podprostor je celé \mathbb{R}^2 , tedy geometrická násobnost λ je také 2. ♦

Příklad 6.3. Matice $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ má jediné vlastní číslo $\lambda = 1$ s algebraickou násobností 2. Jemu příslušný vlastní podprostor je $\text{span}\{(-1, 1)\}$, tedy geometrická násobnost λ je 1. ♦

Věta 6.2. Pro každé vlastní číslo platí, že jeho geometrická násobnost není větší než jeho algebraická násobnost.

Před důkazem věty ukážeme, že charakteristický polynom je invariantní vůči změně báze \mathbb{R}^n , neboli je stejný pro podobné matice. Opravdu, pro každou regulární \mathbf{C} (matice přechodu) platí

$$p_{\mathbf{C}^{-1}\mathbf{A}\mathbf{C}}(\lambda) = \det(\mathbf{C}^{-1}\mathbf{A}\mathbf{C} - \lambda\mathbf{I}) = \det(\mathbf{C}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{C}) = \det(\mathbf{A} - \lambda\mathbf{I}) = p_{\mathbf{A}}(\lambda).$$

Důkaz. (★) Necht' λ je vlastní číslo matice \mathbf{A} a $k \leq n$ jeho geometrická násobnost. Necht' $\mathbf{v}_1, \dots, \mathbf{v}_k$ je báze vlastního podprostoru příslušného λ , tedy $\mathbf{A}\mathbf{v}_i = \lambda\mathbf{v}_i$ pro $i = 1, \dots, k$. Doplníme tuto bázi na bázi $\mathbf{v}_1, \dots, \mathbf{v}_n$ prostoru \mathbb{R}^n . Pak bude

$$\mathbf{A}' = [\mathbf{v}_1 \cdots \mathbf{v}_n]^{-1} \mathbf{A} [\mathbf{v}_1 \cdots \mathbf{v}_n] = [\mathbf{v}_1 \cdots \mathbf{v}_n]^{-1} [\lambda\mathbf{v}_1 \cdots \lambda\mathbf{v}_k \quad \mathbf{A}\mathbf{v}_{k+1} \cdots \mathbf{A}\mathbf{v}_n] = \begin{bmatrix} \lambda\mathbf{I}_k & \mathbf{D} \\ \mathbf{0} & \mathbf{B} \end{bmatrix},$$

kde \mathbf{B}, \mathbf{D} jsou nějaké matice. Vidíme, že přechod od standardní báze k bázi $\mathbf{v}_1, \dots, \mathbf{v}_n$ diagonalizoval prvních k sloupců+řádků matice \mathbf{A} . Nyní (pozor, proměnná polynomu je μ a ne λ)

$$\begin{aligned} p_{\mathbf{A}}(\mu) = p_{\mathbf{A}'}(\mu) &= \det(\mathbf{A}' - \mu\mathbf{I}) = \det \begin{bmatrix} \lambda\mathbf{I}_k - \mu\mathbf{I}_k & \mathbf{D} \\ \mathbf{0} & \mathbf{B} - \mu\mathbf{I}_{n-k} \end{bmatrix} \\ &= \det(\lambda\mathbf{I}_k - \mu\mathbf{I}_k) \det(\mathbf{B} - \mu\mathbf{I}_{n-k}) = (\lambda - \mu)^k p_{\mathbf{B}}(\mu). \end{aligned}$$

Poslední výraz ukazuje, že algebraická násobnost vlastního čísla λ je nejméně k . ■

Z Vět 6.1 a 6.2 plyne, že všechna vlastní čísla, jejich algebraické násobnosti a jim příslušné vlastní podprostory lze reprezentovat seznamem $\lambda_1, \dots, \lambda_n$ vlastních čísel (kde každé číslo je v seznamu tolikrát, kolik je jeho algebraická násobnost) a k nim příslušnými vhodně vybranými vlastními vektory $\mathbf{v}_1, \dots, \mathbf{v}_n$. Např. má-li matice rozměru $n = 3$ dvě různá vlastní čísla s algebraickými násobnostmi 1 a 2, pak je $\lambda_1 \neq \lambda_2 = \lambda_3$, \mathbf{v}_1 je vlastní vektor příslušný vlastnímu číslu λ_1 a $\text{span}\{\mathbf{v}_2, \mathbf{v}_3\}$ je vlastní podprostor příslušný vlastnímu číslu $\lambda_2 = \lambda_3$. Jestliže geometrická násobnost čísla $\lambda_2 = \lambda_3$ je 2, tak vektory $\mathbf{v}_2, \mathbf{v}_3$ jsou lineárně nezávislé (tedy tvoří bázi vlastního podprostoru), jinak jsou lineárně závislé. Soustavu

$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad i = 1, \dots, n \tag{6.6}$$

můžeme napsat jako jedinou maticovou rovnici

$$\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda} \tag{6.7}$$

kde $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$ je diagonální a $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ (roznásobte levou a pravou strany rovnice (6.7) dle pravidla pro násobení blokových matic!).

Lze dokázat (důkaz neuvádíme, i když není obtížný), že vektory $\mathbf{v}_1, \dots, \mathbf{v}_n$ jsou lineárně nezávislé právě tehdy, když geometrická násobnost každého vlastního čísla je rovna jeho algebraické násobnosti. V tom případě je matice \mathbf{V} regulární a (6.7) lze psát jako

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}. \tag{6.8}$$

Vztahu (6.8) se pak říká **rozklad matice podle vlastních čísel** nebo **spektrální rozklad**. Můžeme ho psát také jako $\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathbf{\Lambda}$, což říká, že matice \mathbf{A} je podobná⁴ diagonální matici (neboli **diagonalizovatelná**). Příklad nediagonalizovatelné matice je v Příkladu 6.3.

⁴Zopakujme: matice \mathbf{A} a \mathbf{B} jsou si *podobné*, existuje-li regulární matice \mathbf{C} tak, že $\mathbf{B} = \mathbf{C}\mathbf{A}\mathbf{C}^{-1}$.

Jak se počítají vlastní čísla a vektory? Charakteristický polynom je hlavně teoretický nástroj a přímé hledání jeho kořenů je vhodné jen pro malé matice⁵. Pro větší matice se používají numerické iterační algoritmy, přičemž pro matice různého typu jsou vhodné různé algoritmy. Zásadní rozdíl oproti např. QR rozkladu je v tom, že vlastní čísla obecně nelze spočítat konečným počtem operací sčítání, odčítání, dělení a k -té odmocniny. To má hluboký důvod již v tom, že kořeny polynomu většího než čtvrtého stupně obecně nelze spočítat konečným počtem těchto operací. V Matlabu se algoritmus na vlastní čísla a vektory volá funkcí $[V,D]=\text{eig}(A)$, která vrátí matici Λ (označenou D) a matici V splňující (6.7).

6.1.1 Spektrální rozklad symetrické matice

Pro obecnou čtvercovou matici A mohou být vlastní čísla (a tedy i vlastní vektory) komplexní a matice nemusí být diagonalizovatelná. Příznivá situace nastane, když je matice A symetrická: v tom případě jsou všechna její vlastní čísla reálná a z jejích vlastních vektorů lze vybrat ortonormální množinu n vektorů (toto dokážeme nepovinně níže). Pak tedy v (6.8) je matice Λ reálná a matici V lze zvolit ortogonální ($V^{-1} = V^T$) a platí důležitá věta:

Věta 6.3. Každou symetrickou matici $A \in \mathbb{R}^{n \times n}$ lze rozložit jako

$$A = V\Lambda V^T = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \dots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T \quad (6.9)$$

kde matice $V = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ je ortogonální a $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$.

Zároveň jsme vpravo v (6.9) uvedli i druhou formu rozkladu jako součet dyád (viz §2.2) (roznásobte výraz $V\Lambda V^T$ dle pravidel pro násobení blokových matic!). Vlastní čísla v rozkladu (6.9) je zvykem (matlabská funkce `eig` to tak dělá) řadit vzestupně,

$$\lambda_1 \leq \dots \leq \lambda_n, \quad (6.10)$$

což lze vždy zařídit vhodnou permutací sloupců matice V a diagonálních prvků matice Λ .

Hodnota matice A je rovna hodnotě matice Λ (což plyne např. z Tvzení 3.12, použité dvakrát na výraz $V\Lambda V^T$). Ale hodnota diagonální matice Λ je počet jejích nenulových prvků, tedy nenulových vlastních čísel. Je-li $\text{rank } A = r < n$, pak $n - r$ vlastních čísel je tedy nulových a v rozkladu (6.9) můžeme vynechat jim odpovídající sloupce+řádky matice Λ a sloupce matice V . Tedy je $A = V\Lambda V^T$ pro nějaké $V \in \mathbb{R}^{n \times r}$, $V^T V = I$ a $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$. To odpovídá vynechání $n - r$ nulových sčítanců v sumě dyád v (6.9).

⁵Naopak, hledání kořenů libovolného polynomu lze převést na hledání vlastních čísel matice, která se nazývá *doprovodná matice* (*companion matrix*) polynomu.

Příklad 6.4. Zde je spektrální rozklad symetrické matice 3×3 hodnosti 2:

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

(*) **Důkaz věty o spektrálním rozkladu symetrické matice.** Důkaz je rozdělen na dvě části: Věta 6.4 říká, že symetrická matice má všechna vlastní čísla reálná, a Tvrzení 6.5 příp. Věta 6.6 říkají, že z vlastních vektorů symetrické matice lze vybrat ortonormální množinu n vektorů. Tato část je nepovinná, ale není složitá a je poučné se jí prokousat.

K důkazu reálnosti vlastních čísel budeme potřebovat několik pojmů z komplexní algebry⁶:

- *Komplexně sdružené* číslo ke komplexnímu číslu $x = a + bi \in \mathbb{C}$ je číslo $\bar{x} = a - bi$. Číslo x je reálné, právě když $\bar{x} = x$. Pro matici \mathbf{A} definujeme $\bar{\mathbf{A}}$ po prvcích.
- *Absolutní hodnota* čísla $x \in \mathbb{C}$ je reálné číslo $|x| = (\bar{x}x)^{1/2} = (a^2 + b^2)^{1/2}$.
- *Adjungovaná* matice k matici $\mathbf{A} \in \mathbb{C}^{m \times n}$ je matice $\mathbf{A}^* = \bar{\mathbf{B}}^T \in \mathbb{C}^{n \times m}$. Matice \mathbf{A} je *samoadjungovaná* (neboli *hermitovská*), je-li $\mathbf{A}^* = \mathbf{A}$. Pro reálné matice to znamená $\mathbf{A}^T = \mathbf{A}$.
- *Skalární součin* vektorů $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ je $\mathbf{x}^* \mathbf{y} = \bar{x}_1 y_1 + \dots + \bar{x}_n y_n$ (všimněte si, že $\mathbf{x}^* \mathbf{y} = \overline{\mathbf{y}^* \mathbf{x}}$).
- *Norma* vektoru $\mathbf{x} \in \mathbb{C}^n$ je $\|\mathbf{x}\| = (\mathbf{x}^* \mathbf{x})^{1/2} = (|x_1|^2 + \dots + |x_n|^2)^{1/2}$.

Věta 6.4. Každá hermitovská matice má všechna vlastní čísla reálná.

Důkaz. Necht' $\mathbf{A}^* = \mathbf{A}$ a $\mathbf{A} \mathbf{v} = \lambda \mathbf{v}$. Pak

$$\lambda \|\mathbf{v}\|^2 = \lambda \mathbf{v}^* \mathbf{v} = \mathbf{v}^* \mathbf{A} \mathbf{v} = \mathbf{v}^* \mathbf{A}^* \mathbf{v} = (\mathbf{A} \mathbf{v})^* \mathbf{v} = (\lambda \mathbf{v})^* \mathbf{v} = \bar{\lambda} \mathbf{v}^* \mathbf{v} = \bar{\lambda} \|\mathbf{v}\|^2.$$

Jelikož $\mathbf{v} \neq \mathbf{0}$, je tedy $\lambda = \bar{\lambda}$, tedy $\lambda \in \mathbb{R}$. ■

Existenci ortonormální množiny n vlastních vektorů nejprve dokážeme ve Tvrzení 6.5 pro speciální případ, kdy matice má n různých vlastních čísel: v tomto případě je důkaz velmi krátký. Obecný případ (kdy některá vlastní čísla mohou být násobná), dokazuje Věta 6.5.

Tvrzení 6.5. Necht' $\lambda_i \neq \lambda_j$ jsou dvě různá vlastní čísla symetrické reálné matice \mathbf{A} a $\mathbf{v}_i, \mathbf{v}_j$ jsou jim příslušné vlastní vektory. Pak $\mathbf{v}_i \perp \mathbf{v}_j$.

⁶Je překvapivé a mrzuté, že abychom dokázali, že *reálná* symetrická matice má *reálná* vlastní čísla, potřebujeme k tomu komplexní čísla. Existuje i důkaz používající pouze reálná čísla, ale je mnohem složitější a vyžaduje matematickou analýzu.

Důkaz. Máme

$$\lambda_j \mathbf{v}_i^T \mathbf{v}_j = \mathbf{v}_i^T (\mathbf{A} \mathbf{v}_j) = (\mathbf{A} \mathbf{v}_i)^T \mathbf{v}_j = \lambda_i \mathbf{v}_i^T \mathbf{v}_j.$$

Z toho $(\lambda_i - \lambda_j) \mathbf{v}_i^T \mathbf{v}_j = 0$, z toho $\mathbf{v}_i^T \mathbf{v}_j = 0$. ■

K důkazu poslední věty budeme potřebovat přípravu. I když jsme vlastní čísla a vektory definovali pro matici $\mathbf{A} \in \mathbb{R}^{n \times n}$, lze je obecněji definovat pro lineární transformaci $\mathbf{f}: X \rightarrow X$ na libovolném lineárním podprostoru $X \subseteq \mathbb{R}^n$ podmínkou

$$\mathbf{f}(\mathbf{v}) = \lambda \mathbf{v}. \quad (6.11)$$

Jestliže $X \neq \{0\}$, z argumentu v §6.1 plyne, že \mathbf{f} má aspoň jedno vlastní číslo a vektor⁷.

Podprostor $X \subseteq \mathbb{R}^n$ nazveme *invariantní podprostor* lineární transformace $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ jestliže $\mathbf{f}(X) \subseteq X$, tj. $\mathbf{x} \in X \Rightarrow \mathbf{A}\mathbf{x} \in X$. Tedy \mathbf{f} zobrazí libovolný prvek z X opět do X . V tom případě je restrikce (viz §1.1.2) $\mathbf{f}|_X$ zobrazení \mathbf{f} na množinu X sama o sobě lineární transformace na X , tedy máme $\mathbf{f}|_X: X \rightarrow X$ místo pouhého $\mathbf{f}|_X: X \rightarrow \mathbb{R}^n$. Jestliže je X netriviální, dle základní věty algebry má tedy aspoň jedno vlastní číslo a vektor.

Věta 6.6. *Z vlastních vektorů každé symetrické matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ lze vybrat ortonormální množinu n vektorů.*

Důkaz. Použijeme indukci: ukážeme, že když pro nějaké $k < n$ již máme vlastní vektory $\mathbf{v}_1, \dots, \mathbf{v}_k$, pak existuje vlastní vektor na ně ortogonální.

Dokážeme nejprve, že jestliže \mathbf{A} je symetrická, podprostor $X = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}^\perp$ je invariantní vůči \mathbf{f} . Necht' $\mathbf{x} \in X$, tedy $\mathbf{x}^T \mathbf{v}_1 = \dots = \mathbf{x}^T \mathbf{v}_k = 0$. Máme ukázat, že $\mathbf{A}\mathbf{x} \in X$. To platí, protože $\mathbf{x}^T \mathbf{v}_i = 0 \Rightarrow (\mathbf{A}\mathbf{x})^T \mathbf{v}_i = \mathbf{x}^T \mathbf{A}^T \mathbf{v}_i = \mathbf{x}^T \mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{x}^T \mathbf{v}_i = 0$.

Protože X je invariantní vůči \mathbf{f} , restrikce $\mathbf{f}|_X$ zobrazení \mathbf{f} na množinu X je lineární transformace na X . Protože $k < n$, je $X \neq \{0\}$. Tedy $\mathbf{f}|_X$ má aspoň jeden vlastní vektor. Protože tento vektor patří do X , je ortogonální na $\mathbf{v}_1, \dots, \mathbf{v}_k$.

Takto jsme tedy našli vlastní vektory $\mathbf{v}_1, \dots, \mathbf{v}_n$ splňující $\mathbf{v}_i \perp \mathbf{v}_j$ pro každé $i \neq j$. Když je znormalizujeme, tvoří ortonormální množinu. ■

6.2 Kvadratická forma

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

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j \quad (6.12)$$

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

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j = \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n (a_{ij} + a_{ji}) x_i x_j = \frac{1}{2} \mathbf{x}^T (\mathbf{A} + \mathbf{A}^T) \mathbf{x}. \quad (6.13)$$

⁷Maticově by důkaz byl takhle: Necht' $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, $\mathbf{C} \in \mathbb{R}^{m \times n}$, $X = \text{rng } \mathbf{A} = \text{rng } \mathbf{B}$, $m = \dim X$ a $\mathbf{B}\mathbf{C} = \mathbf{I}$. Jestliže \mathbf{u} je nyní vlastní vektor matice $\mathbf{C}\mathbf{A}\mathbf{B} \in \mathbb{R}^{m \times m}$, pak $\mathbf{v} = \mathbf{B}\mathbf{u}$ je vlastní vektor matice \mathbf{A} . Opravdu, vynásobením rovnice $\mathbf{C}\mathbf{A}\mathbf{B}\mathbf{u} = \lambda \mathbf{u}$ zleva maticí \mathbf{B} dostaneme $\mathbf{A}\mathbf{B}\mathbf{u} = \lambda \mathbf{B}\mathbf{u}$, tedy $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$.

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

Abychom si procvičili maticovou algebru, dokážeme rovnost (6.13) i jinak. Každou čtvercovou matici lze jednoznačně napsat jako součet symetrické a antisymetrické části (viz Cvičení 2.16):

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

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 6.5. 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. ♦

6.2.1 Definitnost kvadratické formy / její matice

Čtvercovou 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, obvykle 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 6.7. *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, funkce f není nikde 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} pozitivně definitní, je forma nulová jen v počátku a všude jinde kladná, tedy počátek je ostré minimum. 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 $\mathbf{y}^T\mathbf{A}\mathbf{y} < 0$. ■

6.2.2 Definitnost ze znamének hlavních minorů

Uvedeme kritérium, které určí definitnost matice podle znamének determinantů jistých jejích podmatic. Pro matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ a neprázdnou množinu $I \subseteq \{1, \dots, n\}$, symbol $\mathbf{A}_I = [a_{ij}]_{i,j \in I} \in \mathbb{R}^{|I| \times |I|}$ označuje matici vytvořenou z prvků matice \mathbf{A} v řádcích a sloupcích s indexy I .

- **Hlavní minor** matice \mathbf{A} je číslo $\det \mathbf{A}_I$ pro nějakou neprázdnou $I \subseteq \{1, \dots, n\}$.
- **Vůdčí hlavní minor** matice \mathbf{A} je její hlavní minor pro nějaké $I = \{1, \dots, k\}$ a $1 \leq k \leq n$.

Počet všech hlavních minorů matice $n \times n$ je $2^n - 1$, počet všech vůdčích hlavních minorů je n .

Příklad 6.6. Příklady matic \mathbf{A}_I pro matici $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{4 \times 4}$. Determinanty první a druhé matice jsou vůdčí hlavní minory matice \mathbf{A} , determinanty třetí a čtvrté jsou pouze hlavní minory.

$$\begin{matrix}
 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} & \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \begin{bmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{bmatrix} & \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \\
 I=\{1,2,3\} & I=\{1,2\} & I=\{1,3,4\} & I=\{2,3\}
 \end{matrix}$$



Věta 6.8. Symetrická matice je

- *positivně definitní, právě když všechny její vůdčí hlavní minory jsou kladné,*
- *positivně semidefinitní, právě když všechny její hlavní minory jsou nezáporné.*

Důkaz neuvádíme. První tvrzení věty je známé **Sylvestrovo kritérium**. Druhé tvrzení lze prakticky použít jen pro malé matice, neboť počet hlavních minorů matice rychle roste s n .

Pozor ovšem: neplatí (a věta to neříká), že symetrická matice je

- *positivně semidefinitní, jestliže všechny její vůdčí hlavní minory jsou nezáporné*
 (např. matice $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ je indefinitní a má všechny vůdčí hlavní minory nezáporné),
- *negativně definitní, jestliže všechny její vůdčí hlavní minory jsou záporné*
 (např. matice $\begin{bmatrix} -2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$ je indefinitní a má všechny vůdčí hlavní minory záporné),
- *negativně semidefinitní, jestliže všechny její hlavní minory jsou nekladné*
 (např. matice $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ je indefinitní a má všechny hlavní minory nekladné).

Přesto věta umožňuje rozhodnout všech pět případů definitnosti s využitím těchto tvrzení:

- Matice \mathbf{A} je negativně [semi]definitní, právě když matice $-\mathbf{A}$ je positivně [semi]definitní.
- Matice je indefinitní, právě když není ani positivně ani negativně semidefinitní.

6.2.3 Diagonalizace kvadratické formy

Definitnost matice lze ovšem také snadno určit pomocí jejích vlastních čísel. Ze spektrálního rozkladu (6.9) máme

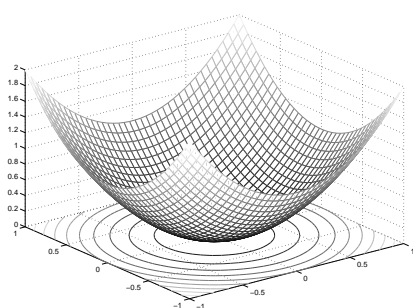
$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 = g(\mathbf{y}) = g(\mathbf{v}^T \mathbf{x}). \quad (6.14)$$

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

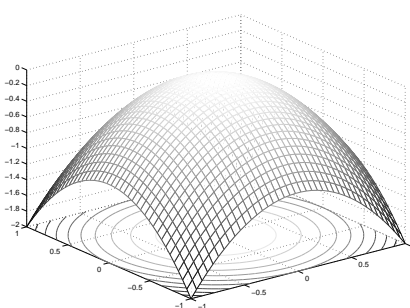
jednoznačná (neboli bijekce, viz §1.1.2). Tedy funkce f a g se liší jen otočením příp. zrcadlením. Totéž můžeme říci i v jazyce bázi: zatímco ve standardní (ortonormální) bázi (tvořené sloupečky jednotkové matice \mathbf{I}) má kvadratická forma matici \mathbf{A} , v ortonormální bázi tvořené sloupci matice \mathbf{V} má forma diagonální matici $\mathbf{\Lambda}$.

Vlastnosti kvadratické formy jsou mnohem lépe patrné z jejího diagonálního tvaru g než z původního f . ‘Kvalitativní’ tvar grafu funkce g je dán vektorem znamének vlastních čísel. Např. pro $n = 2$ si grafy a vrstevnice funkce g (kde bez ztráty obecnosti bereme vrstevnici výšky 1, tj. množinu řešení rovnice $\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = 1$) snadno představíme:

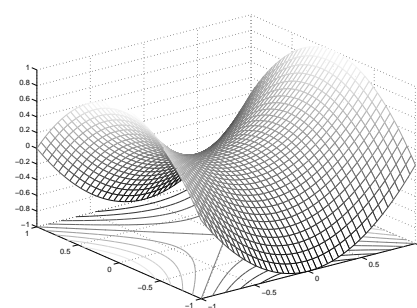
- Je-li $\lambda_1, \lambda_2 > 0$, graf vypadá jako ‘dolík’ a každá vrstevnice výšky 1 je elipsa (střed má v počátku a její hlavní osy jsou souřadnicové osy).
- Je-li $\lambda_1, \lambda_2 < 0$, graf vypadá jako ‘kopec’ a vrstevnice výšky 1 je prázdná množina.
- Je-li $\lambda_1 \lambda_2 < 0$, graf je ‘sedlo’ a vrstevnice výšky 1 je hyperbola.



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

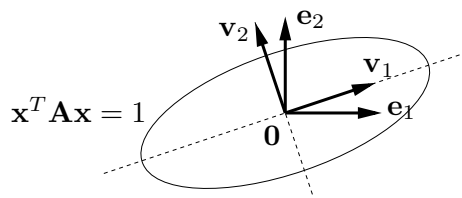


$$g(y_1, y_2) = -y_1^2 - y_2^2$$



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

Vrstevnice výšky 1 obecně (tj. ne nutně diagonální) kvadratické formy (tj. množina řešení $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$) s oběma vlastními čísly kladnými je elipsa se středem v počátku, směry jejíchž hlavních os jsou vlastní vektory:



Věta 6.9. 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 (6.14). Protože transformace $\mathbf{x} = \mathbf{V} \mathbf{y}$ je bijekce, např. $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ platí pro všechna $\mathbf{x} \in \mathbb{R}^n$ právě tehdy, když $\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} \geq 0$ platí pro všechna $\mathbf{y} \in \mathbb{R}^n$. Tedy definitnost matice \mathbf{A} je stejná jako definitnost matice $\mathbf{\Lambda}$. Ale definitnost diagonální matice $\mathbf{\Lambda}$ je okamžitě vidět ze znamének čísel λ_i . Např. výraz (6.14) je nezáporný pro každé \mathbf{y} , právě když všechna λ_i jsou nezáporná. ■

6.2.4 Choleského rozklad

Věta 6.10 (Choleského rozklad). Pro každou symetrickou pozitivně semidefinitní matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ existuje horní trojúhelníková matice $\mathbf{R} \in \mathbb{R}^{n \times n}$ tak, že $\mathbf{A} = \mathbf{R}^T \mathbf{R}$.

Důkaz. Nejprve položíme $\mathbf{B} = \mathbf{\Lambda}^{1/2} \mathbf{V}^T$, kde $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ je spektrální rozklad \mathbf{A} a označili jsme $\mathbf{\Lambda}^{1/2} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Ted' tedy $\mathbf{B}^T \mathbf{B} = \mathbf{V} (\mathbf{\Lambda}^{1/2})^T \mathbf{\Lambda}^{1/2} \mathbf{V} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T = \mathbf{A}$.

Pak uděláme QR rozklad $\mathbf{B} = \mathbf{Q} \mathbf{R}$. Ted' $\mathbf{A} = \mathbf{B}^T \mathbf{B} = \mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} = \mathbf{R}^T \mathbf{R}$. ■

Všimněte si, že když \mathbf{A} není pozitivně semidefinitní, důkaz nefunguje protože $\mathbf{\Lambda}^{1/2}$ má některé diagonální prvky imaginární. Lze dokázat, že pokud je \mathbf{A} pozitivně definitní, tak matice \mathbf{R} je určena jednoznačně.

I když důkaz Věty 6.10 používá spektrální rozklad následovaný QR rozkladem, na výpočet Choleského rozkladu existuje mnohem jednodušší, rychlejší a numericky stabilnější algoritmus. Z Matlabu se tento algoritmus zavolá příkazem `chol` (nastudujte si ho příkazem `help chol`!). Algoritmus umí poznat, když matice není pozitivně semidefinitní.

Typické použití je pro řešení lineární soustavy $\mathbf{Ax} = \mathbf{b}$ se symetrickou pozitivně (semi)definitní maticí. Pro takové soustavy je tato metoda rychlejší a numericky stabilnější než Gaussova eliminace (LU rozklad). Vyskytují se např. v každé iteraci některých numerických iteračních algoritmů (§10). Soustavu $\mathbf{Ax} = \mathbf{R}^T \mathbf{Rx} = \mathbf{b}$ vyřešíme jednoduše tak, že nejdřív spočteme \mathbf{y} ze soustavy $\mathbf{R}^T \mathbf{y} = \mathbf{b}$ zpětnou substitucí a pak \mathbf{x} z $\mathbf{Rx} = \mathbf{y}$ další zpětnou substitucí.

Shrnutím Vět 6.8, 6.9 a 6.10 je užitečná charakterizace pozitivně (semi)definitních matic:

Důsledek 6.11. Pro symetrickou matici \mathbf{A} jsou tato tvrzení ekvivalentní:

- Matice \mathbf{A} je pozitivně definitní.
- Všechny vůdčí hlavní minory matice \mathbf{A} jsou kladné.
- Všechna vlastní čísla matice \mathbf{A} jsou kladná.
- Existuje čtvercová regulární matice \mathbf{B} taková, že $\mathbf{A} = \mathbf{B}^T \mathbf{B}$.

Důsledek 6.12. Pro symetrickou matici \mathbf{A} jsou tato tvrzení ekvivalentní:

- Matice \mathbf{A} je pozitivně semidefinitní.
- Všechny hlavní minory matice \mathbf{A} jsou nezáporné.
- Všechna vlastní čísla matice \mathbf{A} jsou nezáporná.
- Existuje čtvercová nebo úzká matice \mathbf{B} taková, že $\mathbf{A} = \mathbf{B}^T \mathbf{B}$.

6.3 Kvadratická funkce

Kvadratická funkce je polynom (ne nutně homogenní) druhého stupně. Lze jej psát v maticovém tvaru⁸

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{x} + c, \quad (6.15)$$

⁸Jak už jsme zmínili v §2.4 pro afinní zobrazení, kvadratickou funkci nemusíme vždy dostat zadanou přímo ve tvaru (6.15). Příkladem jsou funkce $f(\mathbf{x}, \mathbf{y}) = \mathbf{y}^T (\mathbf{Ax} - \mathbf{b})$ či $f(\mathbf{X}) = \text{tr}((\mathbf{AX} + \mathbf{B})^T (\mathbf{AX} + \mathbf{B}))$. Převést takové funkce do tvaru (6.15) může stát dost práce a umu.

kde $\mathbf{A}^T = \mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$ a $c \in \mathbb{R}$. Oproti⁹ kvadratické formě tedy přibyl lineární a konstantní člen. Všimněte si, že pro $n = 1$ je (6.15) 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.

6.3.1 Doplnění na čtverec

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

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A} (\mathbf{x} - \mathbf{x}_0) + y_0. \quad (6.16)$$

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

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

(kde jsme použili, že $\mathbf{x}^T \mathbf{A} \mathbf{x}_0 = (\mathbf{x}^T \mathbf{A} \mathbf{x}_0)^T = \mathbf{x}_0^T \mathbf{A}^T \mathbf{x} = \mathbf{x}_0^T \mathbf{A} \mathbf{x}$). Porovnáním členů stejného stupně máme

$$\mathbf{b} = -2\mathbf{A} \mathbf{x}_0, \quad (6.17a)$$

$$c = \mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 + y_0, \quad (6.17b)$$

Pokud soustava (6.17a) má řešení, počítáme z ní \mathbf{x}_0 a pak z druhé rovnice y_0 . Pokud soustava (6.17a) nemá řešení, doplnění na čtverec není možné.

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

Příklad 6.7. Máme kvadratickou funkci

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

Její doplnění na čtverec je

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

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

⁹Pro $\mathbf{A} = \mathbf{0}$ 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 6.8. Kvadratická funkce

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

doplňit na čtverec nejde. Funkce tedy nemá na \mathbb{R}^2 extrém – což je ale jasné, protože při konstantním x můžeme změnou y dosáhnout libovolně malé i libovolně velké hodnoty funkce. Načrtněte si graf této funkce! \blacklozenka

Příklad 6.9. Řešme znovu úlohu (5.2). Účelová funkce této úlohy je kvadratická, je totiž

$$\begin{aligned} \|\mathbf{Ax} - \mathbf{b}\|^2 &= (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) \\ &= (\mathbf{x}^T \mathbf{A}^T - \mathbf{b}^T) (\mathbf{Ax} - \mathbf{b}) \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b} \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - 2\mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b}, \end{aligned} \quad (6.18)$$

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 (6.17a) 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 (6.17a)), 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

$$\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} = (\mathbf{Ax})^T \mathbf{Ax} = \|\mathbf{Ax}\|^2 \geq 0. \quad (6.19)$$

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

6.3.2 Kvadrika

Vrstevnice kvadratické funkce se nazývá **kvadrika**. Tedy kvadrika je množina¹⁰

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{x} + c = 0\} \quad (6.20)$$

všech řešení kvadratické rovnice, neboli množina všech kořenů kvadratické funkce.

Když $\mathbf{b} = \mathbf{0}$, množina (6.20) 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 (6.20) je pouhá nadrovina. Když \mathbf{A} nemá plnou hodnost, kvadrika je degenerovaná. Množina (6.20) může být i prázdná.

Jestliže kvadratická funkce dovoluje doplnění na čtverec, můžeme množinu (6.20) psát jako

$$\{\mathbf{x} \in \mathbb{R}^n \mid (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0) + y_0 = 0\}, \quad (6.21)$$

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**¹¹. Když některá vlastní čísla jsou nulová (tedy \mathbf{A} nemá plnou hodnost), kvadrika je degenerovaná. Toto ale nevyčerpává všechny typy

¹⁰Bez ztráty obecnosti uvažujeme vrstevnici výšky nula (tj. množina kořenů) kvadratické funkce (6.15), neboť c můžeme zvolit libovolně.

¹¹Někteří autoři myslí elipsoidem množinu i s vnitřkem, někteří jen její hranici. Rozdíl je stejný jako mezi sférou a koulí.

degenerace: další typy degenerace nastanou, když \mathbf{A} nemá plnou hodnost a funkce nedovoluje doplnění na čtverec.

Předpokládejme, že y_0 je takové, že množina (6.21) obsahuje nekonečný počet bodů. Pro $n = 2$ se pak kvadratika nazývá **kuželosečka** (angl. *conic*). Je-li matice \mathbf{A} pozitivně či negativně definitní maticí, kuželosečka je *elipsa* (speciálně, pokud $\lambda_1 = \lambda_2$ pak je to kružnice), je-li \mathbf{A} indefinitní, je to *hyperbola*.

6.4 Cvičení

6.1. Pro každou z těchto funkcí určete, zda je to polynom. Pokud ano, určete počet proměnných a stupeň polynomu a rozhodněte, jestli je polynom homogenní.

- a) $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = (x^2 + y^2)(x - y) + xy - x - y$
- b) $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \mathbf{a}^T \mathbf{x}$ kde \mathbf{a} je dáno
- c) $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \|\mathbf{x}\|$
- d) $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} + \mathbf{b}\|^2$ kde \mathbf{A}, \mathbf{b} jsou dány
- e) $f: \mathbb{R}^{2n} \rightarrow \mathbb{R}, f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$
- f) $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, f(\mathbf{X}) = \mathbf{a}^T \mathbf{X} \mathbf{b}$ kde \mathbf{a}, \mathbf{b} jsou dány
- g) $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, f(\mathbf{X}) = \det \mathbf{X}$

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

6.3. Napište rovnici, jejímiž kořeny jsou vlastní čísla matice $\begin{bmatrix} 2 & 0 & 3 \\ 0 & -2 & -1 \\ 3 & -1 & 2 \end{bmatrix}$.

6.4. Jaká jsou vlastní čísla a vlastní vektory (a) nulové, (b) jednotkové, (c) diagonální matice? Jaká jsou vlastní čísla trojúhelníkové matice?

6.5. Známe vlastní čísla a vektory matice \mathbf{A} . Jaká jsou vlastní čísla a vektory matice $\mathbf{A} + \alpha \mathbf{I}$?

6.6. 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 vlastních vektorů odpovídajících stejným nenulovým vlastním číslům?

6.7. Každá čtvercová matice splňuje

$$\det \mathbf{A} = \lambda_1 \cdots \lambda_n, \quad (6.22a)$$

$$\operatorname{tr} \mathbf{A} = \lambda_1 + \cdots + \lambda_n. \quad (6.22b)$$

Dokažte, že tyto rovnosti platí pro diagonalizovatelné matice. Pro nediagonalizovatelné matice je dokazovat nemusíte.

6.8. Určete definitnost těchto symetrických matic. Pro všechny matice to udělejte pomocí znamének hlavních minorů, pro matice 2×2 také pomocí vlastních čísel.

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & -2 & 1 \\ -2 & 3 & 0 \\ 1 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} -2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & -2 \end{bmatrix}$$

6.9. Mějme matici $\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 2 & -4 \end{bmatrix}$. Která z následujících tvrzení jsou pravdivá?

- Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nezáporný pro každé $\mathbf{x} \in \mathbb{R}^2$.
- Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nekladný pro každé $\mathbf{x} \in \mathbb{R}^2$.
- Funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ má v bodě $\mathbf{x} = \mathbf{0}$ extrém.

6.10. Máme kvadratickou formu dvou proměnných $f(x, y) = 3x^2 + 2xy + 3y^2$.

- Napište ji ve tvaru $f(x, y) = \begin{bmatrix} x & y \end{bmatrix} \mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix}$ se symetrickou \mathbf{A} .
- Najděte $a, b \in \mathbb{R}$ a ortogonální \mathbf{U} tak, že $f(x, y) = au^2 + bv^2$, kde $\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{U} \begin{bmatrix} x \\ y \end{bmatrix}$.
- Nakreslete množinu bodů (u, v) splňujících $au^2 + bv^2 = 1$.
- Transformujte tuto množinu do souřadnic (x, y) a nakreslete.

6.11. Je množina $\{(x, y) \in \mathbb{R}^2 \mid x^2 - 3xy + y^2 = 1\}$ elipsa nebo hyperbola? Odůvodněte.

6.12. (★) 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).

6.13. Ukažte, že je-li $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ spektrální rozklad symetrické matice \mathbf{A} , platí $\mathbf{A}^n = \mathbf{V} \mathbf{\Lambda}^n \mathbf{V}^T$.

6.14. Následující kvadratické funkce napište ve tvaru (6.15) se symetrickou \mathbf{A} . Pak najděte jejich extrémy a určete typ každého extrému (použijte doplnění na čtverec).

- $f(x, y) = x^2 + 4xy - 2y^2 + 3x - 6y + 5$
- $f(x, y) = x^2 + 2y^2 - 2xy + 2x - y$

6.15. Máme neorientovaný graf (V, E) s množinou vrcholů $V = \{1, \dots, n\}$ a množinou hran¹². $E \subseteq \binom{V}{2}$ 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 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ vzorcem

$$f(\mathbf{x}) = \sum_{\{i,j\} \in E} (x_i - x_j)^2.$$

- 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 $\mathbf{D} = \text{diag}(\mathbf{A}\mathbf{1})$. Co jsou diagonální prvky matice \mathbf{D} v termínech grafu (V, E) ?
- Necht' $\mathbf{B} \in \{-1, 0, 1\}^{n \times m}$ (kde $m = |E|$) je incidenční matice orientovaného grafu vytvořeného tak, že pro každou hranu neorientovaného grafu (V, E) zvolíme orientaci. Ukažte, že $f(\mathbf{x}) = \|\mathbf{B}^T \mathbf{x}\|^2$, tedy $\mathbf{L} = \mathbf{B} \mathbf{B}^T$, kde tyto výrazy nezávisí na zvolených orientacích hran.

¹² $\binom{V}{k}$ značí množinu všech k -prvkových podmnožin množiny V .

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

6.16. Musí mít pozitivně semidefinitní matice na diagonále nezáporné prvky? Odpověď dokažte.

6.17. Dokažte, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ je pozitivně definitní pro každou matici \mathbf{A} a každé $\mu > 0$.

6.18. Dokažte, že symetrická matice je pozitivně definitní, právě když je pozitivně semidefinitní a invertovatelná.

6.19. Dokažte, že každá symetrická pozitivně definitní matice je invertovatelná a její inverze je také pozitivně definitní. Dokažte

- a) s použitím spektrálního rozkladu,
- b) bez použití spektrálního rozkladu.

6.20. Dokažte:

- a) Pro každou matici \mathbf{A} je matice $\mathbf{A}^T \mathbf{A}$ pozitivně semidefinitní.
- b) Matice $\mathbf{A}^T \mathbf{A}$ je pozitivně definitní, právě když \mathbf{A} má lineárně nezávislé sloupce,
- c) Pro každou pozitivně semidefinitní \mathbf{B} existuje \mathbf{A} tak, že $\mathbf{B} = \mathbf{A}^T \mathbf{A}$.
- d) Pro každou pozitivně definitní \mathbf{B} existuje regulární \mathbf{A} tak, že $\mathbf{B} = \mathbf{A}^T \mathbf{A}$.

6.21. (★) 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 $n \times n$.

6.22. (★) Na základě podobnosti relace \preceq z předchozího cvičení a relace \leq na množině \mathbb{R} bychom očekávali, že:

- a) Pokud $\mathbf{A} \preceq \mathbf{B}$ a $\mathbf{C} \preceq \mathbf{D}$, potom $\mathbf{A} + \mathbf{C} \preceq \mathbf{B} + \mathbf{D}$.
- b) Pokud $\mathbf{A} \succeq \mathbf{0}$ a $\alpha \geq 0$, potom $\alpha \mathbf{A} \succeq \mathbf{0}$.
- c) Pokud $\mathbf{A} \succeq \mathbf{0}$, potom $\mathbf{A}^2 \succeq \mathbf{0}$.
- d) Pokud $\mathbf{A} \succeq \mathbf{0}$ a $\mathbf{B} \succeq \mathbf{0}$, potom $\mathbf{AB} \succeq \mathbf{0}$.
- e) 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.

6.23. Co jsou vlastní čísla a vektory projektoru? Odpověď dokažte algebraicky a poté odůvoděte geometrickou úvahou.

6.24. Geometrickou úvahou najděte aspoň dva vlastní vektory a příslušná vlastní čísla reflektoru.

6.25. 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 .

6.26. Dokažte, že každý (ne nutně ortogonální) projektor má vlastní čísla rovna buď 0 nebo 1. Plyne z toho, že každý projektor je pozitivně semidefinitní? Plyne z toho, že každý ortogonální projektor je pozitivně semidefinitní? Odůvodněte.

6.27. Na wikipedii vyhledejte hesla *conic section* (kuželosečka) a *quadric surface* (kvadrika). Všechny typy kuželoseček a kvadrik (nedegenerovaných i denenerovaných) tam uvedené zkuste napsat jako vrstevnici vhodné kvadratické formy nebo funkce.

6.28. (★) Je známo, že elipsa je množina bodů v rovině, které mají od dané dvojice bodů (ohnisek) konstantní součet vzdáleností. Pokud elipsa má střed v počátku, je to tedy množina bodů \mathbf{x} splňujících $\|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} + \mathbf{a}\| = 1$. Člověk by si mohl myslet, že v \mathbb{R}^n bude tato množina elipsoid se středem v počátku, tedy bude popsána rovnicí $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$ pro nějaké pozitivně definitní \mathbf{A} . Je to pravda? Pokud ano, jak se spočítá \mathbf{A} pomocí \mathbf{a} ?

Nápověda a řešení

- 6.5. Vlastní čísla se zvětší o α . Vlastní vektory jsou stejné.
- 6.6. Necht' $\mathbf{A} \mathbf{B} \mathbf{v} = \lambda \mathbf{v}$ kde $\mathbf{v} \neq \mathbf{0}$ (vlastní vektory nesmí být nulové) a $\lambda \neq 0$ (předpoklad). Z toho plyne $\mathbf{B} \mathbf{A} \mathbf{B} \mathbf{v} = \lambda \mathbf{B} \mathbf{v}$ a $\mathbf{B} \mathbf{v} \neq \mathbf{0}$, tedy $\mathbf{B} \mathbf{A} \mathbf{u} = \lambda \mathbf{u}$ kde $\mathbf{u} = \mathbf{B} \mathbf{v} \neq \mathbf{0}$.
- 6.7. Použijte spektrální rozklad (6.8) a vlastnosti determinantu a stopy.
- 6.8. indefinitní, pozitivně definitní, indefinitní, pozitivně semidefinitní, indefinitní, pozitivně definitní, indefinitní
- 6.9. Matice \mathbf{A} není symetrická. Musíme ji nejdříve symetrizovat, tj. vzít matici $\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$, a pak teprv počítat vlastní čísla. Tím zjistíme, že žádné tvrzení neplatí.

6.10.a) $\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

6.10.b) $a = 2, b = 4, \mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$

6.11. Hyperbola, neboť \mathbf{A} má nenulová vlastní čísla opačných znamének.

6.14.a) Převod na tvar (6.15): $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ -6 \end{bmatrix}$, $c = 5$. Doplnění na čtverec existuje (protože \mathbf{A} je regulární). Funkce nemá extrém (protože \mathbf{A} je indefinitní), má sedlo v bodě $(\frac{1}{2}, -1)$.

6.14.b) Má minimum v bodě $-(3, 1)/2$.

6.15.a) Je $(x_i - x_j)^2 = x_i^2 - 2x_i x_j + x_j^2$. Tedy f je homogenní polynom stupně dva, což je kvadratická forma.

6.15.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.

6.15.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.

6.15.d) Prvek d_{ii} matice \mathbf{D} je stupeň (tedy počet incidentních hran) vrcholu i . Je $\mathbf{L} = \mathbf{D} - \mathbf{A}$, neboť

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

6.16. Musí. Stačí vzít $\mathbf{x} = \mathbf{e}_i$ vektory standardní báze.

6.17. $\mathbf{x}^T (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}) \mathbf{x} = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} + \mu \mathbf{x}^T \mathbf{I} \mathbf{x} = \|\mathbf{A} \mathbf{x}\|^2 + \mu \|\mathbf{x}\|^2 > 0$ pro každé $\mathbf{x} \neq \mathbf{0}$.

6.18. 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. Jak jsme řekli v §6.1.1, \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.

6.19.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 $\mathbf{\Lambda}^{-1}$.

6.19.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 3.11 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 $\mathbf{y} \neq \mathbf{0}$.

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

6.20.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 (5.6).

6.20.c) Ve spektrálním rozkladu $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ má $\mathbf{\Lambda}$ nezáporné diagonální prvky. Položme $\mathbf{A} = \mathbf{\Lambda}^{1/2} \mathbf{V}^T$.

6.20.d) Ve spektrálním rozkladu $\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ má $\mathbf{\Lambda}$ kladné diagonální prvky. Položme $\mathbf{A} = \mathbf{\Lambda}^{1/2} \mathbf{V}^T$.

6.21. 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 $\mathbf{x}^T \mathbf{A} \mathbf{x} \leq \mathbf{x}^T \mathbf{B} \mathbf{x} \leq \mathbf{x}^T \mathbf{C} \mathbf{x} \Rightarrow \mathbf{x}^T \mathbf{A} \mathbf{x} \leq \mathbf{x}^T \mathbf{C} \mathbf{x}$.

6.23. Pro $\mathbf{y} \in \text{rng } \mathbf{P}$ platí $\mathbf{P} \mathbf{y} = \mathbf{y}$. Jestliže $\mathbf{P} \neq \mathbf{I}$ (což u projektoru je totéž, jako že je singularní), pak existuje $\mathbf{y} \neq \mathbf{0}$ takové, že $\mathbf{P} \mathbf{y} = \mathbf{0}$. Tedy $\mathbf{P} \neq \mathbf{I}$ má vlastní čísla 0 a 1.

Dokážeme, že \mathbf{P} nemá jiná vlastní čísla než 0 nebo 1. Necht' $\mathbf{P} \mathbf{y} = \lambda \mathbf{y}$, z čehož plyne $\mathbf{P}^2 \mathbf{y} = \mathbf{P} \mathbf{y} = \lambda \mathbf{P} \mathbf{y}$, tj. $\mathbf{x} = \lambda \mathbf{x}$ kde $\mathbf{x} = \mathbf{P} \mathbf{y}$. Z toho plyne, že buď $\mathbf{x} = \mathbf{0}$ a $\lambda = 0$, nebo $\mathbf{x} \neq \mathbf{0}$ a $\lambda = 1$.

Vlastní vektory \mathbf{y} příslušné $\lambda = 0$ splňují $\mathbf{P} \mathbf{y} = \mathbf{0}$, tedy tvoří podprostor null \mathbf{P} bez počátku. To geometricky odpovídá tomu, že všechny vektory z null \mathbf{P} se promítají do počátku. Vlastní vektory příslušné $\lambda = 1$ splňují $\mathbf{P} \mathbf{y} = \mathbf{y}$, což (díky vlastnostem projektoru) platí jen pro $\mathbf{y} \in \text{rng } \mathbf{P}$, tedy tvoří podprostor rng \mathbf{P} bez počátku. To odpovídá tomu, že vektory z rng \mathbf{P} se projekcí nezmění.

6.26. Projektor \mathbf{P} splňuje $\mathbf{P}^2 = \mathbf{P}$. Je-li λ vlastní číslo \mathbf{P} , pak $\mathbf{P} \mathbf{v} = \lambda \mathbf{v}$ pro nějaké $\mathbf{v} \neq \mathbf{0}$. Vynásobením maticí \mathbf{P} zleva z toho máme $\mathbf{P}^2 \mathbf{v} = \mathbf{P} \mathbf{v} = \lambda \mathbf{P} \mathbf{v}$. Jestliže $\mathbf{P} \mathbf{v} \neq \mathbf{0}$, pak $\lambda = 1$. Jestliže $\mathbf{P} \mathbf{v} = \mathbf{0}$, pak $\lambda = 0$ (neboť $\mathbf{v} \neq \mathbf{0}$). Ale pozor, z toho ještě neplyne že matice \mathbf{P} je pozitivně semidefinitní, protože nemusí být symetrická. Pokud je ale \mathbf{P} ortogonální projektor, je to symetrická matice a je tedy pozitivně semidefinitní (geometrický význam: pro každé \mathbf{x} je $\mathbf{x}^T \mathbf{P} \mathbf{x} \geq 0$, tj. úhel mezi vektory \mathbf{x} a $\mathbf{P} \mathbf{x}$ je pravý nebo ostrý, dle (4.3)).

Kapitola 7

PCA a SVD

7.1 Úloha na nejmenší stopu

Spektrální rozklad symetrické matice je mocný nástroj. Ukážeme nyní, že díky němu vlastní čísla a vektory symetrické matice úzce souvisí s jistými jednoduchými optimalizačními úlohami. Na základě těchto teoretických výsledků později formulujeme algoritmus na proložení bodů podprostorem (PCA).

Nechť tedy $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ je spektrální rozklad symetrické matice $\mathbf{A} \in \mathbb{R}^{n \times n}$, tedy matice $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ je diagonální a $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ je ortogonální. Předpokládáme, že vlastní čísla jsou vzestupně seřazena, $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

Věta 7.1. Platí

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

přičemž minimální hodnota se nabývá pro¹ $\mathbf{x} = \mathbf{v}_1$.

Důkaz. Myšlenka důkazu je jednoduchá: nahradíme-li matici \mathbf{A} v úloze (7.1) jejím spektrálním rozkladem, řešení úlohy se stane očividným. Zopakujme rovnost (6.14),

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2,$$

kde jsme substitucí $\mathbf{x} = \mathbf{V} \mathbf{y}$ (neboli $\mathbf{y} = \mathbf{V}^T \mathbf{x}$) uvedli kvadratickou formu do diagonálního tvaru. Protože \mathbf{V} je ortogonální, je $\mathbf{x}^T \mathbf{x} = \mathbf{y}^T \mathbf{V}^T \mathbf{V} \mathbf{y} = \mathbf{y}^T \mathbf{y}$ 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 úloha²

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

Elementární úvahou (viz Příklad 12.4 a Cvičení 7.1) vidíme, že minimum úlohy (7.2) nastane 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 $\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1$. ■

Je snad jasné, že nahradíme-li minimum v úloze (7.1) maximem, získáme místo nejmenšího vlastního čísla λ_1 největší vlastní číslo λ_n . Ted' se asi zeptáte, zda lze formulovat podobné úlohy i pro mezilehlá vlastní čísla $\lambda_2, \dots, \lambda_{n-1}$ a jejich vlastní vektory. Věta 7.2 ukazuje jeden způsob:

¹Netvrdíme, že se minimum nabývá *pouze* v bodě \mathbf{v}_1 . Je-li vlastní číslo λ_1 násobné, nabývá se ve více bodech.

²Převod na úlohu (7.2) vlastně ukazuje, že Větu 7.1 stačí dokázat jen pro diagonální \mathbf{A} .

přidáme-li k úloze (7.1) omezení, že \mathbf{x} je kolmé na vlastní vektory $\mathbf{v}_1, \dots, \mathbf{v}_k$, získáme λ_{k+1} a \mathbf{v}_{k+1} . Kdybychom tedy uměli řešit úlohy typu (7.3), mohli bychom všechna vlastní čísla a vektory matice \mathbf{A} postupně spočítat jedno po druhém.

Věta 7.2. *Necht' $k < n$. Platí*

$$\min\{ \mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1, \mathbf{v}_1^T \mathbf{x} = \dots = \mathbf{v}_k^T \mathbf{x} = 0 \} = \lambda_{k+1} \quad (7.3)$$

a minimální hodnota se nabývá pro $\mathbf{x} = \mathbf{v}_{k+1}$.

Důkaz. Jako v důkazu předchozí věty použijeme spektrální rozklad a dosadíme $\mathbf{x} = \mathbf{V} \mathbf{y}$ (neboli $\mathbf{y} = \mathbf{V}^T \mathbf{x}$). Protože $\mathbf{v}_i = \mathbf{V} \mathbf{e}_i$, je $\mathbf{v}_i^T \mathbf{x} = \mathbf{e}_i^T \mathbf{V}^T \mathbf{x} = \mathbf{e}_i^T \mathbf{y} = y_i$. Tedy podmínky $\mathbf{v}_i^T \mathbf{x} = 0$ se změň na podmínky $\mathbf{e}_i^T \mathbf{y} = y_i = 0$ a úloha se zjednoduší na

$$\min\{ \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 \mid y_1, \dots, y_n \in \mathbb{R}, y_1^2 + \dots + y_n^2 = 1, y_1 = \dots = y_k = 0 \}.$$

Očividně, minimum nastane pro $\mathbf{y} = \mathbf{e}_{k+1}$. To odpovídá $\mathbf{x} = \mathbf{V} \mathbf{e}_{k+1} = \mathbf{v}_{k+1}$ a $\mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_{k+1}$. ■

Úloha (7.3) je zobecněním úlohy (7.1). Úlohu (7.1) lze ale zobecnit ještě jinak:

$$\min \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \dots + \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k \quad (7.4a)$$

$$\text{za podm. } \mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n \text{ jsou ortonormální} \quad (7.4b)$$

Tedy hledáme ortonormální množinu vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ (ortonormální bázi nějakého podprostoru \mathbb{R}^n dimenze k), která minimalizuje součet kvadratických forem (7.4a). Dříve než prozradíme řešení úlohy, zapíšeme ji v jiném značení. Označme $\mathbf{X} = [\mathbf{x}_1 \ \dots \ \mathbf{x}_k] \in \mathbb{R}^{n \times k}$ matici se sloupci $\mathbf{x}_1, \dots, \mathbf{x}_k$. Podmínka na ortonormalitu vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ se dá nyní zapsat jako $\mathbf{X}^T \mathbf{X} = \mathbf{I}$. Dále platí (roznásobte blokové matice a použijte definici stopy!)

$$\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \dots + \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k = \text{tr}([\mathbf{x}_1 \ \dots \ \mathbf{x}_k]^T \mathbf{A} [\mathbf{x}_1 \ \dots \ \mathbf{x}_k]) = \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}). \quad (7.5)$$

Tedy úlohu (7.4) lze napsat jako úlohu (7.6) v této větě:

Věta 7.3. *Necht' $k \leq n$. Platí³*

$$\min\{ \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) \mid \mathbf{X} \in \mathbb{R}^{n \times k}, \mathbf{X}^T \mathbf{X} = \mathbf{I} \} = \lambda_1 + \dots + \lambda_k \quad (7.6)$$

a minimum se nabývá pro $\mathbf{X} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_k]$.

Důkaz. Substitucí $\mathbf{X} = \mathbf{V} \mathbf{Y}$ (neboli $\mathbf{Y} = \mathbf{V}^T \mathbf{X}$) opět převedeme úlohu na jednodušší tvar. Je $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = \text{tr}(\mathbf{X}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{X}) = \text{tr}(\mathbf{Y}^T \mathbf{\Lambda} \mathbf{Y})$. Protože $\mathbf{X}^T \mathbf{X} = \mathbf{Y}^T \mathbf{V}^T \mathbf{V} \mathbf{X} = \mathbf{Y}^T \mathbf{Y}$, podmínka $\mathbf{X}^T \mathbf{X} = \mathbf{I}$ je ekvivalentní podmínce $\mathbf{Y}^T \mathbf{Y} = \mathbf{I}$. Tedy úloha (7.6) má stejnou optimální hodnotu jako úloha

$$\min\{ \text{tr}(\mathbf{Y}^T \mathbf{\Lambda} \mathbf{Y}) \mid \mathbf{Y} \in \mathbb{R}^{n \times k}, \mathbf{Y}^T \mathbf{Y} = \mathbf{I} \}. \quad (7.7)$$

Řešme úlohu (7.7). Z vlastností stopy (viz §2.1.6) máme

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

³Namítnete, že úloha (7.6) není ve tvaru (1.13), protože optimalizujeme přes množinu matic a ne vektorů. To je ale jen záležitost značení, neboť matici rozměru $m \times n$ lze vidět jako vektor o mn složkách.

kde čísla p_{11}, \dots, p_{nn} jsou diagonální prvky matice (ortogonálního projektoru) $\mathbf{P} = \mathbf{Y}\mathbf{Y}^T$. Za podmínky $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$ tato čísla splňují rovnost

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

Dále splňují nerovnosti $0 \leq p_{ii} \leq 1$ pro každé $i = 1, \dots, n$ (viz Cvičení 4.23). Uvažujme úlohu

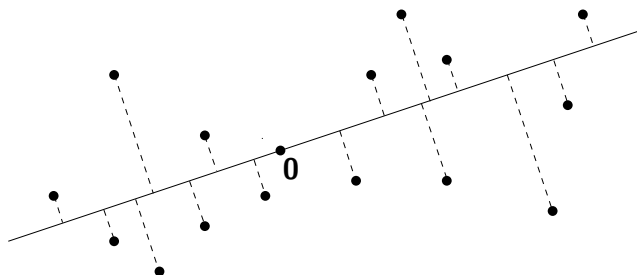
$$\min\{ \lambda_1 p_{11} + \dots + \lambda_n p_{nn} \mid 0 \leq p_{ii} \leq 1 \forall i = 1, \dots, n, p_{11} + \dots + p_{nn} = k \}. \quad (7.8)$$

Tento lineární program vyřešíme úvahou (viz Příklad 12.5): minimální hodnota je $\lambda_1 + \dots + \lambda_k$ a nabývá se pro $p_{11} = \dots = p_{kk} = 1$ a $p_{k+1,k+1} = \dots = p_{nn} = 0$. 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 $\mathbf{X} = \mathbf{V}\mathbf{Y} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_k]$. ■

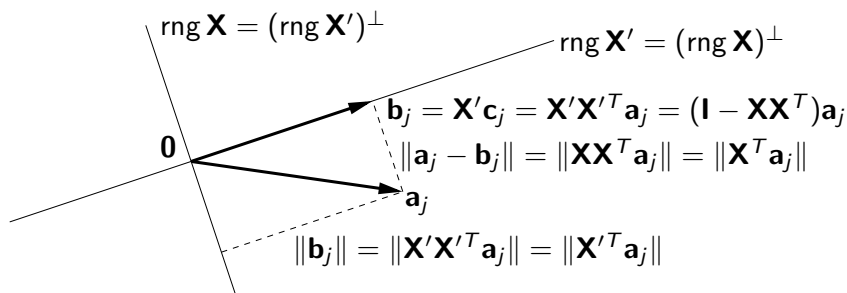
Když nahradíme minimum v úloze (7.6) maximem, analogicky by se dokázalo, že optimální hodnota je $\lambda_{k+1} + \dots + \lambda_n$ a nabývá se pro $\mathbf{X} = [\mathbf{v}_{k+1} \ \dots \ \mathbf{v}_n]$.

7.2 Proložení bodů podprostorem

Necht' jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ a přiřazené číslo $k \leq m$. Hledejme lineární podprostor dimenze k prostoru \mathbb{R}^m , který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_n$. Pro pohodlí uložíme dané body jako sloupce do matice $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n] \in \mathbb{R}^{m \times n}$. Obrázek ukazuje příklad bodů (plná kolečka) a optimálního podprostoru (plná čára) pro $n = 15, m = 2, k = 1$ (minimalizujeme součet druhých mocnin délek přerušovaných úseček):



Abychom jednoduše vyjádřili vzdálenost bodu k podprostoru, budeme místo kýženého podprostoru hledat jeho ortogonální doplněk, reprezentovaný ortonormální bází tvořenou sloupci matice $\mathbf{X} \in \mathbb{R}^{m \times (m-k)}$ (neboť dle Věty 4.6 má tento ortogonální doplněk dimenzi $m - k$). Vzdálenost jednoho bodu \mathbf{a}_j k hledanému podprostoru je délka projekce na jeho ortogonální doplněk (viz §4.6.2), tj. $\|\mathbf{X}\mathbf{X}^T\mathbf{a}_j\| = \|\mathbf{X}^T\mathbf{a}_j\|$. Viz obrázek (je to starý známý obrázek z §4.6 o ortogonální projekci, jen v jiném značení):



Součet čtverců vzdáleností všech bodů k podprostoru je

$$\|\mathbf{X}^T \mathbf{a}_1\|^2 + \dots + \|\mathbf{X}^T \mathbf{a}_n\|^2 = \|\mathbf{X}^T \mathbf{A}\|^2. \quad (7.9)$$

kde $\|\cdot\|$ značí Frobeniovu maticovou normu (2.11). Řešíme tedy úlohu

$$\min\{\|\mathbf{X}^T \mathbf{A}\|^2 \mid \mathbf{X} \in \mathbb{R}^{m \times (m-k)}, \mathbf{X}^T \mathbf{X} = \mathbf{I}\}. \quad (7.10)$$

(Poznamenejme, že úlohu lze formulovat i jako maximalizační, pochopte to z Cvičení 7.14!) Protože $\|\mathbf{X}^T \mathbf{A}\|^2 = \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{A}^T \mathbf{X})$, tuto úlohu vyřešíme pomocí⁴ Věty 7.3. Podrobně platí:

Tvrzení 7.4. Úlohu na proložení bodů podprostorem vyřešíme tak, že spočítáme spektrální rozklad $\mathbf{V} \mathbf{\Lambda} \mathbf{V}^T = \mathbf{A} \mathbf{A}^T$ matice $\mathbf{A} \mathbf{A}^T$ (kde vlastní čísla jsou vzestupně seřazena) a rozdělíme ortogonální matici $\mathbf{V} = [\mathbf{X} \ \mathbf{X}'] \in \mathbb{R}^{m \times m}$ do dvou bloků $\mathbf{X} \in \mathbb{R}^{m \times (m-k)}$ a $\mathbf{X}' \in \mathbb{R}^{m \times k}$. Pak:

- Matice \mathbf{X} je řešením úlohy (7.10) a její sloupce jsou tedy ortonormální báze ortogonálního doplňku hledaného podprostoru.
- Sloupce matice \mathbf{X}' jsou ortonormální báze hledaného podprostoru.

Důkaz. Výrok o matici \mathbf{X} plyne z Věty 7.3. Výrok o matici \mathbf{X}' plyne z Tvrzení 4.10. ■

Příklad 7.1. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_n$ v prostoru \mathbb{R}^3 , jež tvoří sloupce matice $\mathbf{A} \in \mathbb{R}^{3 \times n}$. Necht' $\mathbf{A} \mathbf{A}^T = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ je spektrální rozklad matice $\mathbf{A} \mathbf{A}^T$, kde $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3] \in \mathbb{R}^{3 \times 3}$ 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

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

Číslo $\lambda_1 + \lambda_2$ říká, do jaké míry body neleží v přímce procházející počátkem.

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

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

Číslo λ_1 říká, do jaké míry body neleží v rovině procházející počátkem. ◆

Všimněte si výhody: spočítáme-li jednou spektrální rozklad matice $\mathbf{A} \mathbf{A}^T$, máme rovnou řešení úlohy pro všechna $k = 1, \dots, n$.

Optimální hodnota úlohy (7.10) je chyba proložení našich bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$ podprostorem dimenze k , neboli je to míra, do jaké tyto body neleží v podprostoru dimenze k . Dle Věty 7.3 je tato optimální hodnota rovna

$$\lambda_1 + \dots + \lambda_{m-k}, \quad (7.11)$$

kde $\lambda_1 \leq \dots \leq \lambda_m$ je spektrum matice $\mathbf{A} \mathbf{A}^T$.

⁴Snad je jasné, že \mathbf{A} v §7.2 označuje jinou matici než v §7.1.

7.2.1 Jiný pohled: nejbližší matice nižší hodnosti

Úlohu na optimální proložení bodů podprostorem lze formulovat i jiným, ekvivalentním způsobem (viz obrázek výše): k bodům $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ hledáme body $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathbb{R}^m$, které leží v (neznámém) podprostoru dimenze k a minimalizují součet čtverců vzdáleností

$$\|\mathbf{a}_1 - \mathbf{b}_1\|^2 + \dots + \|\mathbf{a}_n - \mathbf{b}_n\|^2 = \|\mathbf{A} - \mathbf{B}\|^2. \quad (7.12)$$

Výraz (7.12) jsme napsali i v maticové formě, kde $\mathbf{B} = [\mathbf{b}_1 \ \dots \ \mathbf{b}_n] \in \mathbb{R}^{m \times n}$ značí matici se sloupci $\mathbf{b}_1, \dots, \mathbf{b}_n$. Body $\mathbf{b}_1, \dots, \mathbf{b}_n$ leží v podprostoru dimenze k , právě když

$$\dim \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_n\} = \dim \text{rng } \mathbf{B} = \text{rank } \mathbf{B} \leq k. \quad (7.13)$$

Všimněte si, že je tam $\text{rank } \mathbf{B} \leq k$ a ne $\text{rank } \mathbf{B} = k$ (jak byste asi čekali), protože když množina bodů leží v podprostoru dimenze menším než k , pak leží i v podprostoru dimenze k (např. když nějaké body v \mathbb{R}^3 leží na přímce, tak leží také v rovině). Úlohu lze tedy napsat jako⁵

$$\min\{\|\mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{B} \in \mathbb{R}^{m \times n}, \text{rank } \mathbf{B} \leq k\}. \quad (7.14)$$

To ale znamená, že k dané matici \mathbf{A} hledáme nejbližší (ve smyslu Frobeniovy normy) matici \mathbf{B} hodnosti nejvýše k . Tato úloha je angl. známa jako *low rank approximation*.

Tvrzení 7.5. Optimální řešení úlohy (7.14) je

$$\mathbf{B} = \mathbf{X}'\mathbf{X}'^T\mathbf{A} = (\mathbf{I} - \mathbf{X}\mathbf{X}^T)\mathbf{A} \quad (7.15)$$

kde \mathbf{X}, \mathbf{X}' jsou matice z Tvrzení 7.4. Úlohy (7.14) a (7.10) mají stejnou optimální hodnotu.

Důkaz. Z Věty 4.11 plyne, že optimální body $\mathbf{b}_1, \dots, \mathbf{b}_n$ jsou ortogonálními projekcemi bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$ na optimální podprostor, spočítaný v Tvrzení 7.4. Jinak by totiž účelová funkce úlohy (7.14) šla zmenšit posunutím některého bodu v tomto podprostoru.

Ortogonální projektor na tento optimální podprostor je $\mathbf{X}'\mathbf{X}'^T$ a na jeho ortogonální doplněk $\mathbf{X}\mathbf{X}^T = \mathbf{I} - \mathbf{X}'\mathbf{X}'^T$ (viz Tvrzení 4.10). Vzdálenost bodu \mathbf{a}_j od optimálního podprostoru je $\|\mathbf{a}_j - \mathbf{b}_j\| = \|\mathbf{a}_j - \mathbf{X}'\mathbf{X}'^T\mathbf{a}_j\| = \|(\mathbf{I} - \mathbf{X}'\mathbf{X}'^T)\mathbf{a}_j\| = \|\mathbf{X}\mathbf{X}^T\mathbf{a}_j\| = \|\mathbf{X}^T\mathbf{a}_j\|$. Sečtením přes všechna j dostaneme účelové funkce úloh (7.14) a (7.10). ■

Projekci $\mathbf{b} = \mathbf{X}'\mathbf{X}'^T\mathbf{a}$ bodu \mathbf{a} na optimální podprostor můžeme psát jako $\mathbf{b} = \mathbf{X}'\mathbf{c}$, kde $\mathbf{c} \in \mathbf{X}'^T\mathbf{a} \in \mathbb{R}^k$ jsou souřadnice bodu \mathbf{b} v ortonormální bázi tvořené sloupečky matice \mathbf{X}' (viz §4.6). Tedy (7.15) lze psát jako $\mathbf{B} = \mathbf{X}'\mathbf{C}$, kde sloupce $\mathbf{c}_1, \dots, \mathbf{c}_n$ matice $\mathbf{C} = \mathbf{X}'\mathbf{A} \in \mathbb{R}^{k \times n}$ jsou souřadnice bodů $\mathbf{b}_1, \dots, \mathbf{b}_n$ v této bázi. Pro $n \gg m$ řešení úlohy (7.14) můžeme vidět jako kompresi dat uložených v matici \mathbf{A} : body $\mathbf{a}_1, \dots, \mathbf{a}_n$ chceme nahradit body $\mathbf{c}_1, \dots, \mathbf{c}_n$ (které zaberou v paměti daleko méně čísel, přičemž velikost báze \mathbf{X}' je zanedbatelná) tak, aby chyba aproximace (7.12) byla co nejmenší. Srovnej s Větou 3.8 o rank factorization!

7.2.2 Co když prokládáme afinním podprostorem?

Změňme nyní úlohu tak, že místo (lineárního) podprostoru hledáme *afinní* podprostor dimenze k , který minimalizuje součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_n$.

⁵Pozor, v úloze (7.14) je neznámá matice značena jako \mathbf{B} , na což asi nejste zvyklí.

Věta 7.6. Necht' $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ a $k \leq m$. Afinní podprostor dimenze k minimalizující součet čtverců vzdáleností k bodům $\mathbf{a}_1, \dots, \mathbf{a}_n$ prochází jejich těžištěm $\bar{\mathbf{a}} = \frac{1}{n}(\mathbf{a}_1 + \dots + \mathbf{a}_n)$.

Důkaz. Hledaný afinní podprostor parametrizujeme jako (viz Věta 3.18)

$$\{\mathbf{b} \in \mathbb{R}^m \mid \mathbf{X}^T \mathbf{b} = \mathbf{y}\}, \quad (7.16)$$

kde $\mathbf{X} \in \mathbb{R}^{m \times (m-k)}$ je matice s ortonormálními sloupci a $\mathbf{y} \in \mathbb{R}^n$. Vzdálenost bodu $\mathbf{a} \in \mathbb{R}^m$ od tohoto podprostoru je $\|\mathbf{X}^T \mathbf{a} - \mathbf{y}\|$ (viz §4.6.2). Součet čtverců vzdáleností všech bodů k podprostoru je

$$\|\mathbf{X}^T \mathbf{a}_1 - \mathbf{y}\|^2 + \dots + \|\mathbf{X}^T \mathbf{a}_n - \mathbf{y}\|^2.$$

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} pevné a minimalizujeme pouze přes \mathbf{y} , minimum se nabývá v bodě $\mathbf{y} = \mathbf{X}^T \bar{\mathbf{a}}$ (viz Příklad 1.12). Z toho plyne (proč?), že podprostor (7.16) obsahuje bod $\bar{\mathbf{a}}$. ■

Nyní je řešení jasné: nejprve posuneme body $\mathbf{a}_1, \dots, \mathbf{a}_n$ tak, aby jejich těžiště leželo v počátku, a potom najdeme (lineární) podprostor minimalizující součet čtverců vzdáleností k posunutým bodům dle Tvrzení 7.4. Jeho posunutím zpět pak získáme kýžený afinní podprostor.

PCA. Úloha na proložení bodů afinním podprostorem dimenze k je statistice známá jako *rozvoj podle hlavních komponent* (angl. *principal component analysis, PCA*) nebo *Karhunen-Loewův rozvoj*. Podotkněme, že úlohu nelze nijak převést na přibližné řešení lineární soustavy ve smyslu nejmenších čtverců, tedy úlohu (5.2). Ve statistice je matice $\mathbf{A}\mathbf{A}^T = \mathbf{a}_1\mathbf{a}_1^T + \dots + \mathbf{a}_n\mathbf{a}_n^T$ interpretována jako empirická kovarianční matice vzorku $\mathbf{a}_1, \dots, \mathbf{a}_n$ m -rozměrné náhodné veličiny (viz poznámka u Věty 5.2).

7.3 Přeuročené homogenní lineární soustavy

Řešme homogenní lineární soustavu

$$\mathbf{A}\mathbf{x} = \mathbf{0}, \quad (7.17)$$

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 (3.23). Může být homogenní soustava 'přeuročena'? Přeuročenost není rozumné definovat jako u nehomogenní soustavy (viz §5.1), protože homogenní soustava má vždy triviální řešení $\mathbf{x} = \mathbf{0}$. Ovšem, kdybychom 'přeuročenou' soustavu (7.17) 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

$$\min\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\}. \quad (7.18)$$

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 (proložení bodů kuželosečkou). Kuželosečka (viz §6.3.2) je množina

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

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, \dots, f , která minimalizují součet

$$Q(x_1, y_1)^2 + \dots + Q(x_m, y_m)^2. \quad (7.19)$$

Tato formulace ale nevyjadřuje to, co chceme, protože minimum výrazu (7.19) se nabývá pro $a = b = c = d = e = f = 0$. V tomto případě množina K není křivka, ale celá rovina \mathbb{R}^2 , což rozhodně nechceme. Ve skutečnosti navíc potřebujeme, aby aspoň jedno z čísel a, \dots, f bylo nenulové. Toho dosáhneme uvalením podmínky

$$a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1. \quad (7.20)$$

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 (7.19) za podmínky (7.20) se dá psát jako minimalizace $\|\mathbf{A}\boldsymbol{\theta}\|^2$ za podmínky $\boldsymbol{\theta}^T\boldsymbol{\theta} = 1$, kde

$$\mathbf{A} = \begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^2 & x_my_m & y_m^2 & x_m & y_m & 1 \end{bmatrix}, \quad \boldsymbol{\theta} = (a, b, c, d, e, f) \quad \blacklozen$$

Obecněji, soustavu (7.17) nazveme přeuročenu 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 (7.14). Úloha (7.14) odpovídá tomu, že co nejméně změním matici \mathbf{A} , aby prostor řešení soustavy (7.17) měl kýženu dimenzi $n-k$. Neboli nejprve najdeme matici \mathbf{B} s hodnotí k nejbližší matici \mathbf{A} a potom řešíme soustavu $\mathbf{B}\mathbf{x} = \mathbf{0}$, jejíž prostor řešení již má dimenzi $n-k$.

Poznámka: 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 (SVD)

Věta 7.7 (SVD). Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ lze rozložit jako

$$\mathbf{A} = \mathbf{USV}^T = s_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + s_p \mathbf{u}_p \mathbf{v}_p^T \quad (7.21)$$

kde $p = \min\{m, n\}$, matice $\mathbf{S} \in \mathbb{R}^{m \times n}$ je diagonální s diagonálními prvky s_1, \dots, s_p a matice $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_m] \in \mathbb{R}^{m \times m}$ a $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ jsou ortogonální.

Rozklad (7.21) se nazývá **singulární rozklad** (*singular value decomposition, SVD*) matice \mathbf{A} . Diagonální prvky s_1, \dots, s_p matice \mathbf{S} se nazývají **singulární čísla** matice \mathbf{A} . Připomeňme, že diagonální matice nemusí být čtvercová (viz §2), diagonální matice rozměru $m \times n$ má $p = \min\{m, n\}$ diagonálních prvků. Singulární čísla je zvykem (i Matlab to tak dělá) volit nezáporná a řadit je sestupně⁶,

$$s_1 \geq \cdots \geq s_p \geq 0, \quad (7.22)$$

což lze vždy zajistit případným vynásobením čísla s_i a jednoho z vektorů $\mathbf{u}_i, \mathbf{v}_i$ minus jedničkou a permutací sloupců \mathbf{U}, \mathbf{V} a diagonálních prvků \mathbf{S} . Sloupce $\mathbf{u}_1, \dots, \mathbf{u}_p$ matice \mathbf{U} nazývají levé **singulární vektory** a sloupce $\mathbf{v}_1, \dots, \mathbf{v}_p$ matice \mathbf{V} pravé singulární vektory matice \mathbf{A} .

Rozlišujeme tři verze SVD:

- Rozklad $\mathbf{A} = \mathbf{USV}^T$ podle Věty 7.7 je *plné* SVD.
- Pokud je \mathbf{A} obdélníková, pak posledních $m - p$ sloupců \mathbf{U} příp. $n - p$ sloupců \mathbf{V} je v součinu \mathbf{USV}^T násobeno nulami. Jejich vynecháním a oříznutím matice \mathbf{S} získáme *redukované* SVD, kde $\mathbf{S} = \text{diag}(s_1, \dots, s_p) \in \mathbb{R}^{p \times p}$ je diagonální a $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_p] \in \mathbb{R}^{m \times p}$ a $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_p] \in \mathbb{R}^{n \times p}$ mají ortonormální sloupce.
- Platí $\text{rank } \mathbf{A} = \text{rank } \mathbf{S}$ (plyne z Tvzení 3.12), tedy hodnota matice je rovna počtu jejích nenulových singulárních čísel. Pokud $\text{rank } \mathbf{A} = r < p$ (tedy \mathbf{A} nemá plnou hodnotu), pak $s_{r+1} = \cdots = s_p = 0$ a můžeme tedy matice dále zmenšit na $\mathbf{S} = \text{diag}(s_1, \dots, s_r) \in \mathbb{R}^{r \times r}$, $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_r] \in \mathbb{R}^{m \times r}$ a $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_r] \in \mathbb{R}^{n \times r}$. Toto je *kompaktní* verze SVD.

Rozdíl mezi třemi verzemi SVD vzniká jen díky maticovému značení. Napíšeme-li SVD jako sumu dyád v (7.21), rozdíl mezi třemi verzemi se stane triviální: sčítance $s_i \mathbf{u}_i \mathbf{v}_i^T$ pro $i > p$ v této sumě nejsou vůbec, a když je $\text{rank } \mathbf{A} < p$ pak je posledních $p - r$ sčítanců nulových.

Příklad 7.3. Zde je příklad plného a redukovaného (zde i kompaktního) SVD matice 2×3 :

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

Singulární rozklad matice \mathbf{A} má úzký vztah ke spektrálnímu rozkladu matic $\mathbf{A}^T \mathbf{A}$ a $\mathbf{A} \mathbf{A}^T$. Předpokládejme, že rozklad (7.21) existuje. Pak

$$\mathbf{A}^T \mathbf{A} = \mathbf{VS}^T \mathbf{U}^T \mathbf{USV}^T = \mathbf{VS}^T \mathbf{SV}^T, \quad (7.23a)$$

$$\mathbf{A} \mathbf{A}^T = \mathbf{USV}^T \mathbf{VS}^T \mathbf{U}^T = \mathbf{USS}^T \mathbf{U}^T. \quad (7.23b)$$

⁶Na rozdíl od vlastních čísel symetrické matice, která je zvykem řadit vzestupně, dle (6.10).

Ale to jsou spektrální rozklady (6.9) symetrických pozitivně semidefinitních matic $\mathbf{A}^T\mathbf{A}$ a $\mathbf{A}\mathbf{A}^T$. Matice $\mathbf{S}^T\mathbf{S} \in \mathbb{R}^{n \times n}$ a $\mathbf{S}\mathbf{S}^T \in \mathbb{R}^{m \times m}$ jsou diagonální a jejich nenulové diagonální prvky jsou druhé mocniny nenulových diagonálních prvků matice \mathbf{S} . 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 6.6). Právě a levé singulární vektory jsou vlastní vektory těchto matic příslušné nenulovým vlastním číslům.

Ovšem to ještě není důkaz Věty 7.7, protože matice získané nezávisle dvěma spektrálními rozklady (7.23a) a (7.23b) nemusí splňovat (7.21). Důkaz je zde:

Důkaz ()*. Stačí dokázat existenci kompaktního SVD, tedy najít diagonální $\mathbf{S} \in \mathbb{R}^{r \times r}$ a $\mathbf{U} \in \mathbb{R}^{m \times r}$ a $\mathbf{V} \in \mathbb{R}^{n \times r}$ tak, že $\mathbf{U}^T\mathbf{U} = \mathbf{I} = \mathbf{V}^T\mathbf{V}$ a $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$.

Matice $\mathbf{A}^T\mathbf{A}$ je symetrická pozitivně semidefinitní (Cvičení 6.20), tedy má nezáporná vlastní čísla. Dle (5.6) je $\text{rank}(\mathbf{A}^T\mathbf{A}) = \text{rank} \mathbf{A} = r$, tedy matice $\mathbf{A}^T\mathbf{A}$ má r nenulových vlastních čísel. Její spektrální rozklad můžeme psát jako $\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{S}^2\mathbf{V}^T$, kde jsme z matice \mathbf{S}^2 odstranili sloupce+řádky a z matice \mathbf{V} sloupce odpovídající nulovým vlastním číslům, takže $\mathbf{S}^2 \in \mathbb{R}^{r \times r}$ má na diagonále kladná čísla (tudíž je regulární) a $\mathbf{V} \in \mathbb{R}^{n \times r}$ má ortonormální sloupce. Položme

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

Matice $\mathbf{U}, \mathbf{S}, \mathbf{V}$ tvoří kompaktní SVD matice \mathbf{A} , neboť

$$\mathbf{U}^T\mathbf{U} = \mathbf{S}^{-1}\mathbf{V}^T\mathbf{A}^T\mathbf{A}\mathbf{V}\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{V}^T\mathbf{V}\mathbf{S}^2\mathbf{V}^T\mathbf{V}\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{S}^2\mathbf{S}^{-1} = \mathbf{I}, \quad (7.24a)$$

$$\mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{A}\mathbf{V}\mathbf{S}^{-1}\mathbf{S}\mathbf{V}^T = \mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{A}. \quad (7.24b)$$

Jediné, co zde není očividné, je rovnost $\mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{A}$, neboť sice $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ ale už ne nutně $\mathbf{V}\mathbf{V}^T = \mathbf{I}$ (protože matice \mathbf{V} není čtvercová). Dokažme tuto rovnost. Platí

$$\text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{A}^T\mathbf{A}) = \text{rng}(\mathbf{V}\mathbf{S}^2\mathbf{V}^T) = \text{rng} \mathbf{V}, \quad (7.25)$$

kde první rovnost plyne z (5.5a) a třetí rovnost z Tvzení 3.12 (protože matice $\mathbf{S}^2\mathbf{V}^T$ má lineárně nezávislé řádky). Matice $\mathbf{V}\mathbf{V}^T$ je ortogonální projektor na podprostor $\text{rng} \mathbf{V} = \text{rng}(\mathbf{A}^T)$ a proto pro každé $\mathbf{x} \in \text{rng}(\mathbf{A}^T)$ platí $\mathbf{V}\mathbf{V}^T\mathbf{x} = \mathbf{x}$. Protože každý sloupec matice \mathbf{A}^T patří do $\text{rng}(\mathbf{A}^T)$, máme $\mathbf{V}\mathbf{V}^T\mathbf{A}^T = \mathbf{A}^T$ neboli $\mathbf{A}\mathbf{V}\mathbf{V}^T = \mathbf{A}$. ■

Důkaz ukazuje, že SVD lze spočítat ze spektrálního rozkladu matice $\mathbf{A}^T\mathbf{A}$ (nebo alternativně $\mathbf{A}\mathbf{A}^T$). Tento způsob ale není numericky vhodný, protože výpočet součinu $\mathbf{A}^T\mathbf{A}$ či $\mathbf{A}\mathbf{A}^T$ může vést ke zbytečným zaokrouhlovacím chybám (viz §5.1.2). Na SVD byly proto vymyšleny algoritmy, které se explicitnímu výpočtu těchto součinů vyhýbají. Tedy kdykoliv musíte počítat spektrální rozklad matice tvaru $\mathbf{A}^T\mathbf{A}$ nebo $\mathbf{A}\mathbf{A}^T$, je správné použít algoritmus na SVD. Matlabský příkaz `[U,S,V]=svd(A)` počítá plné SVD a `[U,S,V]=svd(A,'econ')` počítá redukované SVD. Minimální verzi SVD Matlab nenabízí, protože není dobrý nápad, aby v nějakém algoritmu rozměry výstupních matic závisely na hodnotě vstupní matice.

7.4.1 Nejblíže matice nižší hodnosti z SVD

V §7.2.1 jsme odvodili řešení úlohy (7.14) pomocí spektrálního rozkladu matice $\mathbf{A}\mathbf{A}^T$. Následující klasická věta formuluje toto řešení elegantně pomocí SVD matice \mathbf{A} .

Věta 7.8 (Eckart-Young). Necht' $\mathbf{A} = \mathbf{USV}^T$ je SVD matice \mathbf{A} a $k \leq p = \min\{m, n\}$. Řešení úlohy (7.14) je

$$\mathbf{B} = \mathbf{US}_k \mathbf{V}^T = s_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + s_k \mathbf{u}_k \mathbf{v}_k^T \quad \text{kde} \quad \mathbf{S}_k = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{S} \quad (7.26)$$

(tedy matice \mathbf{S}_k se získá vynulováním $p - k$ nejmenších diagonálních prvků matice \mathbf{S}).

Důkaz. Důkaz plyne ze vztahu singulárního a spektrálního rozkladu a z Věty 7.5. Dle (7.23b) můžeme matici \mathbf{U} ve spektrálním rozkladu $\mathbf{AA}^T = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ získat jako matici \mathbf{U} v singulárním rozkladu $\mathbf{A} = \mathbf{USV}^T$. Protože však vlastní čísla jsou řazená sestupně ale singulární čísla sestupně, blok této matice odpovídající k největším vlastním číslům matice \mathbf{AA}^T je $\mathbf{X} \in \mathbb{R}^{m \times k}$ kde $\mathbf{U} = [\mathbf{X} \quad \mathbf{X}']$. Dle Věty 7.5 je tedy optimální řešení úlohy (7.14) rovno $\mathbf{B} = \mathbf{XX}^T \mathbf{A}$, což rozepíšeme jako $\mathbf{B} = \mathbf{XC}$ a $\mathbf{C} = \mathbf{X}^T \mathbf{A}$. Zbytek je cvičení na úpravy maticových výrazů:

$$\mathbf{C} = \mathbf{X}^T \mathbf{A} = \mathbf{X}^T \mathbf{USV}^T = \mathbf{X}^T [\mathbf{X} \quad \mathbf{X}'] \mathbf{SV}^T = [\mathbf{X}^T \mathbf{X} \quad \mathbf{X}^T \mathbf{X}'] \mathbf{SV}^T = [\mathbf{I}_k \quad \mathbf{0}] \mathbf{SV}^T$$

$$\mathbf{B} = \mathbf{XC} = \mathbf{X} [\mathbf{I}_k \quad \mathbf{0}] \mathbf{SV}^T = \mathbf{X} [\mathbf{X} \quad \mathbf{X}'] \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{SV}^T = \mathbf{U} \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{SV}^T = \mathbf{US}_k \mathbf{V}^T.$$

Rozměry nulových matic se liší podle toho, zda je SVD plné nebo redukované (věta platí pro obě verze SVD). ■

Všimněte si, že suma dyád v (7.26) je vlastně zkrácená suma v (7.21).

7.4.2 Extrémy lineární funkce isometrie

Zde ukážeme, jak optimalizovat lineární funkci matice s ortonormálními sloupci (reprezentující tedy isometrii). Lineární funkce matice $\mathbf{X} \in \mathbb{R}^{m \times n}$ má obecný tvar $f(\mathbf{X}) = \langle \mathbf{A}, \mathbf{X} \rangle = \sum_{i,j} a_{ij} x_{ij}$ (viz §2.1.7), kde $\mathbf{A} \in \mathbb{R}^{m \times n}$. Tato úloha (7.27) je jednodušší než (7.6), je to nejjednodušší smysluplná úloha s omezením na ortonormalitu (protože co může být jednoduššího než lineární funkce?).

Věta 7.9. Necht' $m \geq n$ a $\mathbf{A} \in \mathbb{R}^{m \times n}$. Úloha

$$\max\{ \langle \mathbf{A}, \mathbf{X} \rangle \mid \mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{X}^T \mathbf{X} = \mathbf{I} \} \quad (7.27)$$

má optimální řešení $\mathbf{X} = \mathbf{UV}^T$ kde $\mathbf{A} = \mathbf{USV}^T$ je redukované SVD matice \mathbf{A} .

Důkaz. Uvažujme nejdříve plné SVD $\mathbf{A} = \mathbf{USV}^T$. Platí $\langle \mathbf{A}, \mathbf{X} \rangle = \langle \mathbf{USV}^T, \mathbf{X} \rangle = \langle \mathbf{S}, \mathbf{U}^T \mathbf{XV} \rangle$ (viz §2.1.7). Zavedeme substituci $\mathbf{Y} = \mathbf{U}^T \mathbf{XV}$. Protože \mathbf{U} a \mathbf{V} jsou ortogonální, $\mathbf{X}^T \mathbf{X} = \mathbf{I}$ platí právě když $\mathbf{Y}^T \mathbf{Y} = \mathbf{I}$. Tedy úloha (7.27) je ekvivalentní úloze

$$\max\{ \langle \mathbf{S}, \mathbf{Y} \rangle \mid \mathbf{Y} \in \mathbb{R}^{m \times n}, \mathbf{Y}^T \mathbf{Y} = \mathbf{I} \}. \quad (7.28)$$

Protože $m \geq n$ a \mathbf{S} je diagonální, máme $\langle \mathbf{S}, \mathbf{Y} \rangle = s_1 y_{11} + \cdots + s_n y_{nn}$. Protože $\mathbf{Y}^T \mathbf{Y} = \mathbf{I}$, prvky matice \mathbf{Y} splňují nerovnosti $-1 \leq y_{ij} \leq 1$ (viz Cvičení 4.22). Protože $s_1, \dots, s_n \geq 0$, výraz $\langle \mathbf{S}, \mathbf{Y} \rangle$ proto nemůže být větší než $s_1 + \cdots + s_n$ a tuto hodnotu nabývá pro $y_{11} = \cdots = y_{nn} = 1$. Existuje právě jedna matice $\mathbf{Y} \in \mathbb{R}^{m \times n}$ splňující $y_{11} = \cdots = y_{nn} = 1$ a $\mathbf{Y}^T \mathbf{Y} = \mathbf{I}$, a to $\mathbf{Y} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \end{bmatrix}$. To odpovídá $\mathbf{X} = \mathbf{UYV}^T = \mathbf{U} \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T$. Ale matice $\mathbf{U} \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \end{bmatrix}$ je tvořená prvními n sloupci matice \mathbf{U} , tedy $\mathbf{X} = \mathbf{UV}^T$ kde $\mathbf{A} = \mathbf{USV}^T$ je redukované SVD matice \mathbf{A} . ■

Nejbližší isometrie. Tento výsledek lze použít pro nalezení matice s ortonormálními sloupci nejbližší dané matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ s $m \geq n$, tedy pro řešení úlohy

$$\min\{\|\mathbf{X} - \mathbf{A}\|^2 \mid \mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{X}^T \mathbf{X} = \mathbf{I}\}. \quad (7.29)$$

Máme

$$\|\mathbf{X} - \mathbf{A}\|^2 = \langle \mathbf{X} - \mathbf{A}, \mathbf{X} - \mathbf{A} \rangle = \langle \mathbf{X}, \mathbf{X} \rangle - \langle \mathbf{X}, \mathbf{A} \rangle - \langle \mathbf{A}, \mathbf{X} \rangle + \langle \mathbf{A}, \mathbf{A} \rangle = \|\mathbf{X}\|^2 - 2\langle \mathbf{A}, \mathbf{X} \rangle + \|\mathbf{A}\|^2.$$

Ale výrazy $\|\mathbf{A}\|^2$ a $\|\mathbf{X}\|^2 = \text{tr}(\mathbf{X}^T \mathbf{X}) = \text{tr} \mathbf{I} = n$ nezávisí na \mathbf{X} , tedy minimalizace $\|\mathbf{X} - \mathbf{A}\|^2$ je totéž jako maximalizace $\langle \mathbf{A}, \mathbf{X} \rangle$.

Ortogonální Prokrustův problém. Dále lze výsledek použít na úlohu⁷, často potkávanou např. v robotice: máme body $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{R}^n$ a $\mathbf{b}_1, \dots, \mathbf{b}_k \in \mathbb{R}^m$, tvořící sloupce matic $\mathbf{A} \in \mathbb{R}^{n \times k}$ a $\mathbf{B} \in \mathbb{R}^{m \times k}$, kde $m \geq n$. Hledáme takovou isometrickou transformaci \mathbf{X} bodů $\mathbf{a}_1, \dots, \mathbf{a}_k$, aby byly co nejbližší bodům $\mathbf{b}_1, \dots, \mathbf{b}_k$ ve smyslu nejmenších čtverců, tedy aby výraz

$$\sum_{i=1}^k \|\mathbf{X} \mathbf{a}_i - \mathbf{b}_i\|^2 = \|\mathbf{X} \mathbf{A} - \mathbf{B}\|^2$$

byl minimální. Řešíme tedy úlohu

$$\min\{\|\mathbf{X} \mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{X}^T \mathbf{X} = \mathbf{I}\}. \quad (7.30)$$

Opět máme

$$\|\mathbf{X} \mathbf{A} - \mathbf{B}\|^2 = \langle \mathbf{X} \mathbf{A} - \mathbf{B}, \mathbf{X} \mathbf{A} - \mathbf{B} \rangle = \|\mathbf{X} \mathbf{A}\|^2 - 2\langle \mathbf{X} \mathbf{A}, \mathbf{B} \rangle + \|\mathbf{B}\|^2.$$

Ale výrazy $\|\mathbf{X} \mathbf{A}\|^2 = \|\mathbf{A}\|^2$ (viz (4.9)) a $\|\mathbf{B}\|^2$ jsou konstanty, tedy minimalizace $\|\mathbf{X} \mathbf{A} - \mathbf{B}\|^2$ je totéž jako maximalizace $\langle \mathbf{X} \mathbf{A}, \mathbf{B} \rangle = \langle \mathbf{X}, \mathbf{B} \mathbf{A}^T \rangle = \langle \mathbf{B} \mathbf{A}^T, \mathbf{X} \rangle$. Tedy jsme úlohu převedli na tvar (7.27).

Pokud dovolíme body $\mathbf{a}_1, \dots, \mathbf{a}_k$ nejen otočit/ozrcadlit ale i posunout, lze udělat podobnou úvahu jako v §7.2.2. Tak odvodíme, že optimální posunutí je rozdíl těžišť bodů \mathbf{b}_i a \mathbf{a}_i .

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*⁸. 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. Uvažujme výraz $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})$ v úloze (7.6).

⁷Procrústés (neboli napínač) byl loupežník z řecké mytologie, který měl nepříjemný zvyk. Zbloudilého poutníka donutil ulehnout ve své lesní chatě na postel. Když byl nešťastník kratší než postel, natáhl ho na délku postele, a když přečuhoval, tak ho zase od nohou kousek usekl.

⁸'Kvocient' znamená 'podíl' nebo 'zlomek'.

- a) Dokažte, že je-li $X \subseteq \mathbb{R}^n$ podprostor, pak výraz je stejný pro všechny matice \mathbf{X} splňující $X = \text{rng } \mathbf{X}$ a $\mathbf{X}^T \mathbf{X} = \mathbf{I}$, tj. nezávisí na tom jakou ortonormální bázi podprostoru zvolíme.
- b) (*) Chceme najít obecnější výraz, který bude stejný pro všechny matice $\mathbf{X} \in \mathbb{R}^{n \times k}$ splňující $X = \text{rng } \mathbf{X}$ a $\text{rank } \mathbf{X} = k$ (tedy báze už nemusí být ortonormální). Dokažte, že takovým výrazem je $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1})$. Dokažte, že úloha (7.6) má stejnou optimální hodnotu jako úloha $\min \{ \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}) \mid \mathbf{X} \in \mathbb{R}^{n \times k}, \text{rank } \mathbf{X} = k \}$.

7.5. Označme jako $P_k^m = \{ \mathbf{X} \mathbf{X}^T \mid \mathbf{X} \in \mathbb{R}^{m \times k}, \mathbf{X}^T \mathbf{X} = \mathbf{I} \} \subseteq \mathbb{R}^{m \times m}$ množinu všech možných ortogonálních projektorů v \mathbb{R}^m na podprostor dimenze k . Ukažte, že úloha (7.10) je ekvivalentní úlohám $\min_{\mathbf{P} \in P_{m-k}^m} \|\mathbf{P} \mathbf{A}\|$ a $\max_{\mathbf{P} \in P_k^m} \|\mathbf{P} \mathbf{A}\|$.

7.6. 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 podprostor $X \subseteq \mathbb{R}^3$, který minimalizuje součet čtverců kolmých vzdáleností bodů k podprostoru X , kde X je

- přímka procházející počátkem,
- rovina procházející počátkem,
- přímka která může ale nemusí procházet počátkem.

Vždy najděte vektor \mathbf{x}_0 a ortonormální bázi lineárního podprostoru X' tak, aby $X = \mathbf{x}_0 + X'$. Dále spočítejte ortogonální projekce bodů na podprostor X (je snad jasné, co myslíme projekcí na afinní podprostor) a souřadnice těchto projekcí v ortonormální bázi podprostoru X' . Použijte (a) spektrální rozklad, (b) SVD. Můžete použít počítač.

7.7. (*) Necht' v rozkladu (7.21) je $\sigma_1 \geq \dots \geq \sigma_p$ a necht' $\text{rank } \mathbf{A} = r$. Dokažte, že prvních r sloupců matice \mathbf{U} je ortonormální báze prostoru $\text{rng } \mathbf{A}$ a prvních r sloupců matice \mathbf{V} je ortonormální báze prostoru $\text{rng}(\mathbf{A}^T)$.

7.8. (*) V Matlabu se báze nulového prostoru matice najde funkcí `null`; vypište si implementaci této funkce matlabským příkazem `edit null` a najděte souvislost se Cvičením 7.7.

7.9. Je dán spektrální rozklad dané symetrické matice. Jak byste z něho jednoduše našli SVD této matice?

7.10. 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. Pokud byste počítali ručně, napovíme že $\mathbf{U} = \frac{1}{5} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}$.

7.11. 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? Nápověda: použijte výsledek Cvičení 6.6.

7.12. Jsou-li $\mathbf{x}_1, \dots, \mathbf{x}_k$ sloupce matice \mathbf{X} , dokažte tyto rovnosti (srov. (7.5)):

$$\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = \langle \mathbf{A} \mathbf{X}, \mathbf{X} \rangle = \langle \mathbf{A}, \mathbf{X} \mathbf{X}^T \rangle = \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \dots + \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k.$$

7.13. 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.

7.14. Místo úlohy z §7.2, ve které hledáme podprostor dimenze k tak aby součet čtverců bodů $\mathbf{a}_1, \dots, \mathbf{a}_k$ k němu byl minimální, uvažujme úlohu, ve které hledáme podprostor dimenze $m - k$ tak aby součet čtverců bodů k němu byl *maximální*. Jak tyto úlohy spolu souvisejí? Jak z řešení jedné úlohy získám řešení druhé?

7.15. Komutuje operace ortogonální projekce s operací těžiště? Tj., když body $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ ortogonálně promítnu na podprostor $X \subseteq \mathbb{R}^m$ a potom najdu těžiště výsledných bodů, dostanu stejný bod jako když nejdříve spočítám těžiště bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$ a pak ho promítnu na podprostor X ? Odpověď dokažte.

7.16. 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 změním omezení (7.20) 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?

7.17. (*) 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 .

7.18. Někdo by mohl namítnout, že Věta 7.3 plyne okamžitě z Věty 7.2, neboť minimalizaci součtu $\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + \dots + \mathbf{x}_k^T \mathbf{A} \mathbf{x}_k$ za podmínky $\mathbf{x}_i^T \mathbf{x}_j = \delta_{ij}$ jde provést takto: nejdříve minimalizujeme $\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1$ za podmínky $\mathbf{x}_1^T \mathbf{x}_1 = 1$, což má řešení $\mathbf{x}_1 = \mathbf{v}_1$, pak \mathbf{x}_1 zafixujeme a minimalizujeme $\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2$ za podmínky $\mathbf{x}_2^T \mathbf{x}_2 = 1$ a $\mathbf{x}_2^T \mathbf{x}_1 = 0$, což má řešení $\mathbf{x}_2 = \mathbf{v}_2$, atd.

Takovému typu algoritmu se říká ‘hladový’ (angl. *greedy*). Obecný hladový algoritmus je popsán takto. Minimalizujeme účelovou funkci $f_1(x_1) + \dots + f_k(x_k)$ za podmínky $(x_1, \dots, x_k) \in X$, kde X je daná množina přípustných řešení. Algoritmus v první iteraci minimalizuje $f_1(x_1)$ za podmínky $(x_1, \dots, x_k) \in X$. Ve druhé iteraci zafixuje proměnnou x_1 a minimalizuje $f_2(x_2)$ za podmínky $(x_1, \dots, x_k) \in X$. Ve třetí zafixuje x_1, x_2 a minimalizuje $f_3(x_3)$ za podmínky $(x_1, \dots, x_k) \in X$, atd.

Hladový algoritmus obecně nemusí najít globální optimum (i když v případě úlohy (7.6) ho najde). Najděte příklad (tedy konkrétní k , X a f_1, \dots, f_k), kdy se tak stane.

7.19. Napišme plné SVD matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r jako

$$\mathbf{A} = [\mathbf{U} \quad \mathbf{U}'] \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{V} \quad \mathbf{V}']^T = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad (7.31)$$

kde $\mathbf{U} \in \mathbb{R}^{m \times r}$, $\mathbf{U}' \in \mathbb{R}^{m \times (m-r)}$, $\mathbf{V} \in \mathbb{R}^{n \times r}$, $\mathbf{V}' \in \mathbb{R}^{n \times (n-r)}$ a $\mathbf{S} \in \mathbb{R}^{r \times r}$ je regulární. Dokažte, že sloupce matic $\mathbf{U}, \mathbf{U}', \mathbf{V}, \mathbf{V}'$ jsou ortonormální báze čtyř základních podprostorů generovaných maticí \mathbf{A} , tedy podprostorů $\text{rng } \mathbf{A}$, $\text{null}(\mathbf{A}^T)$, $\text{rng}(\mathbf{A}^T)$, $\text{null } \mathbf{A}$.

7.20. Dokažte, že pro libovolnou matici \mathbf{A} existuje limita (5.27) a je rovna pseudoinverzi \mathbf{A}^+ matice \mathbf{A} (dle Věty 5.5). Postupujte takto:

- Dokažte, že $\mathbf{A}_\mu^+ = \mathbf{V}\mathbf{S}_\mu^+\mathbf{U}^T$, kde \mathbf{A}_μ^+ je definované vzorcem (5.18), $\mathbf{S}_\mu^+ = (\mathbf{S}^2 + \mu\mathbf{I})^{-1}\mathbf{S}$ a $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ je kompaktní SVD matice \mathbf{A} .
- Dokažte, že $\mathbf{S}^+ = \lim_{\mu \rightarrow 0^+} \mathbf{S}_\mu^+ = \mathbf{S}^{-1}$.

7.21. 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!

Nápověda a řešení

- 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 $y_1^2 + \dots + y_n^2 = 1$.
- 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.
- b) Z cykličnosti stopy je $\text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}) = \text{tr}(\mathbf{A} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T)$. Matice $\mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ je ortogonální projektor na podprostor \mathbf{X} . Ve Cvičení 5.11 jsme dokázali, že tento projektor nezávisí na změně báze.
- Nápověda, ukažte, že $\|\mathbf{X}^T \mathbf{A}\| = \|\mathbf{X} \mathbf{X}^T \mathbf{A}\| = \|\mathbf{P} \mathbf{A}\|$.
- a) $\mathbf{x}_0 = \mathbf{0}$, ortonorm. báze X' je např. $\{(0.6912, -0.2181, 0.6889)\}$.
- b) $\mathbf{x}_0 = \mathbf{0}$, ortonorm. báze X' je např. $\{(0.6912, -0.2181, 0.6889), (-0.3771, -0.9221, 0.0864)\}$.
- c) $\mathbf{x}_0 = (1.25, -1.25, 0.25)$, ortonorm. báze X' je např. $\{(0.6599, 0.0280, 0.7508)\}$.
- První úloha jde napsat jako (7.10), druhá úloha jako $\max\{\|\mathbf{X}'^T \mathbf{A}\|^2 \mid \mathbf{X}' \in \mathbb{R}^{m \times k}, \mathbf{X}'^T \mathbf{X}' = \mathbf{I}\}$. Optimální řešení jedné úlohy je vždy druhá matice v ortogonální matici $[\mathbf{X} \ \mathbf{X}']$, tedy jsou to matice na obrázku v §7.2 a v Tvzení 7.4. Optimální hodnota obou úloh se liší jen o konstantu $\|\mathbf{A}\|^2$, protože platí Pythagorova věta $\|\mathbf{X}^T \mathbf{a}_j\|^2 + \|\mathbf{X}'^T \mathbf{a}_j\|^2 = \|\mathbf{a}_j\|^2$, v součtu přes všechny body $\|\mathbf{X}^T \mathbf{A}\|^2 + \|\mathbf{X}'^T \mathbf{A}\|^2 = \|\mathbf{A}\|^2$ (viz obrázek).
- Ano, komutují. Nápověda k důkazu: Označíme-li $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n] \in \mathbb{R}^{m \times n}$, pak těžiště bodů je bod $\mathbf{A} \mathbf{t}$ kde $\mathbf{t} = \frac{1}{n} \mathbf{1}_n$ a projekce bodů je $\mathbf{P} \mathbf{A}$ kde \mathbf{P} je projektor. Důkaz dokončete sami.
- Dosaďte $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ (\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 $\mathbf{y} = \mathbf{V} \bar{\mathbf{y}} = \mathbf{v}_1$.
- $k = 2$, $X = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$ je kružnice, $f_1(x) = f_2(x) = x$. Globální minimum je $x_1 = x_2 = -1/\sqrt{2}$ (nakreslete si obrázek). Hladový algoritmus ovšem v prvním kroku minimalizuje x_1 za podmínky $x_1^2 + x_2^2 = 1$, což dá $x_1 = -1$. To už zůstane zafixované.
- Použijte Tvzení 3.12 a 4.10.

Část II

Nelineární optimalizace

Kapitola 8

Derivace

8.1 Nelineární zobrazení

Dosud jsme potkali lineární a afinní zobrazení a kvadratické funkce. Nyní přidáme libovolné diferencovatelné funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (zopakujte funkce a zobrazení z §1.1.3). Pro ty, kdo prošli vícerozměrnou analýzou, bude většina této kapitoly opakování, i když možná v trochu jiném značení.

Většinu definic a vět vyslovíme pro funkce a zobrazení, jejichž definičním oborem je celé \mathbb{R}^n . To samozřejmě obecně neplatí, protože součástí definice funkce je i její definiční obor. Např. funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ daná předpisem $f(x) = x^2$ a funkce $g: [0, 1] \rightarrow \mathbb{R}$ daná předpisem $g(x) = x^2$ jsou dvě různé funkce. Podobně, funkce daná předpisem $f(x) = \sqrt{1 - x^2}$ není pro některá $x \in \mathbb{R}$ definována a jejím největším možným definičním oborem je interval $[-1, 1]$. Toto zjednodušení však usnadní výklad a vždy bude očividné, jak by se látka zobecnila pro jiný definiční obor.

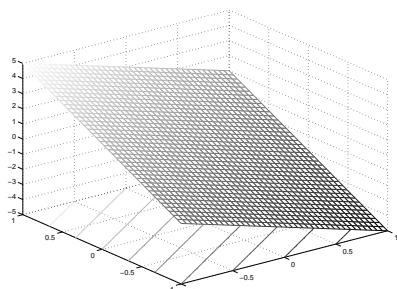
Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ užíváme tyto pojmy:

- **Graf** funkce f je množina¹ $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) = y\} = \{(\mathbf{x}, f(\mathbf{x})) \mid \mathbf{x} \in \mathbb{R}^n\}$.
- **Vrstevnice** funkce f výšky y je množina $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = y\}$.

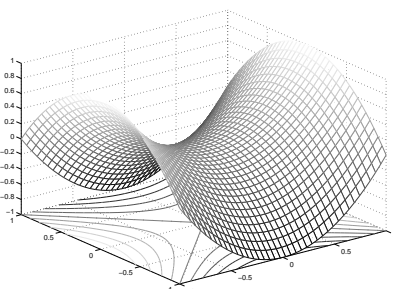
Příklad 8.1. Graf funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ dané vzorcem $f(x) = x^2$ je množina $\{(x, y) \mid x^2 = y\}$. Vrstevnice funkce f výšky 1 je množina $\{x \mid x^2 = 1\} = \{-1, 1\}$.

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 $\{(x, y) \mid x^2 + y^2 = 1\}$. ♦

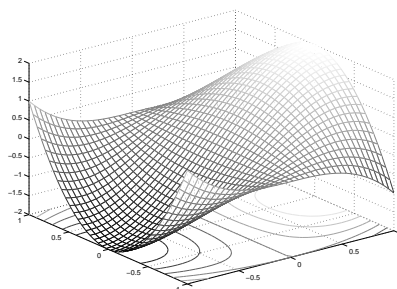
Příklad 8.2. Obrázek ukazuje část (na obdélníku $[-1, 1]^2$) grafu a vrtevníc funkcí $f: \mathbb{R}^2 \rightarrow \mathbb{R}$:



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



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



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

¹Graf funkce je tedy relace (1.1), pomocí které se pojem ‘funkce’ definuje.

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. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$ kde $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{b} \in \mathbb{R}^n, c \in \mathbb{R}$ (kvadratická funkce)
8. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = e^{-\mathbf{x}^T \mathbf{x}}$ (nenormalizované Gaussovo rozdělení)
9. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \max_{i=1}^n x_i$
10. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^2, \mathbf{f}(t) = (\cos t, \sin t)$ (parametrizace kružnice, množina $\mathbf{f}([0, 2\pi))$ je kružnice)
11. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^3, \mathbf{f}(t) = (\cos t, \sin t, at)$ (parametrizace šroubovice neboli helixu)
12. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^n, \mathbf{f}(t) = \mathbf{x} + t\mathbf{v}$ (parametrizace přímky jdoucí bodem \mathbf{x} ve směry \mathbf{v})
13. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n, \mathbf{f}(\mathbf{x}) = \mathbf{x}$ (identické zobrazení neboli identita)
14. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}(\mathbf{x}) = \mathbf{a}$ kde $\mathbf{a} \in \mathbb{R}^m$ (konstantní zobrazení)
15. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ kde $\mathbf{A} \in \mathbb{R}^{m \times n}$
(lineární zobrazení; je to také parametrizace lineárního podprostoru $\mathbf{f}(\mathbb{R}^n) = \text{rng } \mathbf{A} \subseteq \mathbb{R}^m$)
16. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ kde $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$
(afinní zobrazení; je to také parametrizace afinního podprostoru $\mathbf{f}(\mathbb{R}^n) = \mathbf{b} + \text{rng } \mathbf{A} \subseteq \mathbb{R}^m$)
17. $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \mathbf{f}(u, v) = ((R + r \cos v) \cos u, (R + r \cos v) \sin u, r \sin v)$
(parametrizace toru neboli anuloidu, množina $\mathbf{f}([0, 2\pi) \times [0, 2\pi))$ je torus)
18. *Skalární pole* (pojem z fyziky) je funkce $\mathbb{R}^3 \rightarrow \mathbb{R}$. Např. teplotní pole přiřadí každému bodu prostoru teplotu.
19. *Vektorové pole* (opět z fyziky) je zobrazení $\mathbb{R}^3 \rightarrow \mathbb{R}^3$. Např. elektrické pole přiřadí každému bodu v \mathbb{R}^3 vektor z \mathbb{R}^3 .
20. Při technice *image morphing* se obrázek např. obličejů spojitě zdeformuje na obrázek jiného obličejů. Považujeme-li obrázek pro jednoduchost za prostor \mathbb{R}^2 (ve skutečnosti je to množina $\{1, \dots, m\} \times \{1, \dots, n\}$), morphing je tedy zobrazení $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. ♦

8.2 Spojitost

Spojitosť zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ lze definovat několika ekvivalentními způsoby. Uvedeme tento: zobrazení \mathbf{f} je **spojité** v bodě $\mathbf{x} \in \mathbb{R}^n$, jestliže

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{x}). \quad (8.1)$$

Tato definice není pohodlná, protože počítat limity funkcí více proměnných je obtížné. Uvedeme proto postačující (avšak nikoliv nutnou) podmínku pro spojitost, která v praxi většinou stačí. Přitom předpokládáme, že čtenář pouze dokáže ověřit spojitost funkcí jedné proměnné. Věta vlastně ukazuje, jak lze rekurzivně sestavit spojitá zobrazení více proměnných ze spojitých funkcí jedné proměnné. Důkaz věty vynecháme.

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

$$f(x, y) = \sin(x + y^2) \quad (8.2)$$

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. ♦

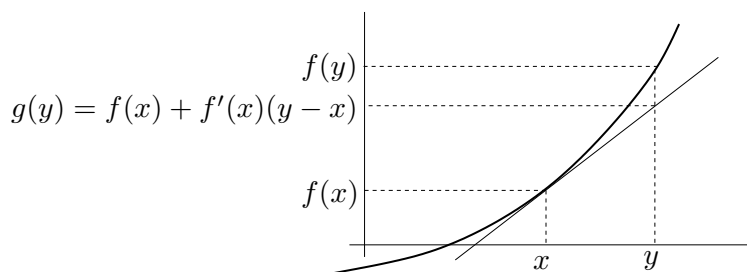
8.3 Derivace funkce jedné proměnné

Zopakujme pojmy diferencovatelnosti a derivace funkce jedné proměnné. Funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ je *diferencovatelná* v bodě $x \in \mathbb{R}$ tehdy, existuje-li limita

$$\lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x} = \lim_{\alpha \rightarrow 0} \frac{f(x + \alpha) - f(x)}{\alpha}. \quad (8.3)$$

Pak se hodnota limity (8.3) nazývá *derivace* funkce f v bodě x . Tuto derivaci značíme jedním ze symbolů $\frac{df(x)}{dx}$ (Leibnizovo značení) nebo $f'(x)$ (Lagrangeovo značení). Symbol $\frac{df}{dx}: \mathbb{R} \rightarrow \mathbb{R}$ nebo $f': \mathbb{R} \rightarrow \mathbb{R}$ označuje funkci, která bodu $x \in \mathbb{R}$ přiřadí derivaci funkce f v bodě x .

Této derivaci se přesněji říká *oboustranná derivace*. *Jednostranná* diferencovatelnost a derivace funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v bodě $x \in \mathbb{R}$ se definuje obdobně, pouze limita (8.3) se změní na limitu zprava či zleva. Značíme je $f'_+(x)$ (derivace zprava) a $f'_-(x)$ (derivace zleva). Oboustranná derivace existuje právě tehdy, když existují derivace zprava i zleva a jsou si rovny.



Diferencovatelnost a derivaci lze definovat i jiným, ekvivalentním způsobem: funkce f je diferencovatelná v bodě x tehdy, lze-li ji v okolí bodu x 'dobře' aproximovat afinní funkcí.

Upřesněme tento výrok. Označme naši afinní funkci jako $g: \mathbb{R} \rightarrow \mathbb{R}$ a pišme ji ve tvaru²

$$g(y) = f(x) + a(y - x), \quad (8.4)$$

Samozřejmě, funkce g je obecně jiná³ pro každé $x \in \mathbb{R}$. Funkce g je opravdu afinní (viz (3.10)), protože $g(y) = ay + b$ kde $b = f(x) - ax$. Tedy funkce f je diferencovatelná v bodě x právě tehdy, když existuje číslo a takové, že chyba aproximace $f(y) - g(y)$ se pro $y \rightarrow x$ v prvním řádu blíží nule, tedy

$$\lim_{y \rightarrow x} \frac{f(y) - g(y)}{|y - x|} = \lim_{y \rightarrow x} \frac{f(y) - f(x) - a(y - x)}{|y - x|} = 0. \quad (8.5)$$

Pak se číslu a říká *derivace* funkce f v bodě x . Dokážeme, že obě definice jsou ekvivalentní:

Tvrzení 8.2. *Limita (8.3) existuje a je rovna a , právě když platí (8.5).*

Důkaz. Oboustranná limita existuje právě když existují obě jednostranné limity. Je

$$\begin{aligned} \lim_{y \rightarrow x^+} \frac{f(y) - f(x) - a(y - x)}{|y - x|} &= \lim_{y \rightarrow x^+} \frac{f(y) - f(x) - a(y - x)}{y - x} = \lim_{y \rightarrow x^+} \frac{f(y) - f(x)}{y - x} - a, \\ \lim_{y \rightarrow x^-} \frac{f(y) - f(x) - a(y - x)}{|y - x|} &= \lim_{y \rightarrow x^-} \frac{f(y) - f(x) - a(y - x)}{-(y - x)} = a - \lim_{y \rightarrow x^-} \frac{f(y) - f(x)}{y - x}. \end{aligned}$$

Porovnáme-li pravé strany s (8.3), dostaneme dokazované tvrzení. ■

8.4 Parciální derivace

Parciální derivaci funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ podle i -té proměnné v bodě $\mathbf{x} = (x_1, \dots, x_n)$ spočítáme tak, že všechny proměnné x_j , $j \neq i$, považujeme za konstanty a zderivujeme funkci podle jediné proměnné x_i . Značíme ji jedním ze symbolů $\frac{\partial f(\mathbf{x})}{\partial x_i}$ nebo $f_{x_i}(\mathbf{x})$. Symbol $\frac{\partial f}{\partial x_i}: \mathbb{R}^n \rightarrow \mathbb{R}$ nebo $f_{x_i}: \mathbb{R}^n \rightarrow \mathbb{R}$ pak označuje funkci, která přiřazuje bodu $\mathbf{x} \in \mathbb{R}^n$ parciální derivaci funkce f podle i -té proměnné v bodě \mathbf{x} .

Příklad 8.5. Parciální derivace funkce (8.2) jsou

$$\frac{\partial f(x, y)}{\partial x} = f_x(x, y) = \cos(x + y^2), \quad (8.6a)$$

$$\frac{\partial f(x, y)}{\partial y} = f_y(x, y) = 2y \cos(x + y^2). \quad (8.6b)$$

◆

²Někdo raději píše $g(x + \alpha) = f(x) + a\alpha$, což je samozřejmě stejné jako (8.4) dosazením $y = x + \alpha$.

³Takže bychom přesněji mohli psát $g_x(y)$ nebo $g(x, y)$.

8.5 Totální derivace

Jak se dají pojmy diferencovatelnosti a derivace zobecnit z funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ na zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$? Mohli bychom zkusit toho dosáhnout zobecněním limity (8.3)⁴. Obvyklejší způsob ovšem vede přes alternativní definici diferencovatelnosti (8.5). Zkusme zobrazení v blízkosti daného bodu $\mathbf{x} \in \mathbb{R}^n$ aproximovat afinním zobrazením

$$\mathbf{g}(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{A}(\mathbf{y} - \mathbf{x}), \quad (8.7)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$. Zobrazení \mathbf{f} se nazývá **totálně diferencovatelné** (nebo jen **diferencovatelné**) v bodě \mathbf{x} , jestliže existuje matice \mathbf{A} taková, že chyba aproximace $\mathbf{f}(\mathbf{y}) - \mathbf{g}(\mathbf{y})$ se pro $\mathbf{y} \rightarrow \mathbf{x}$ v prvním řádu blíží nule, tedy

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} \frac{\mathbf{f}(\mathbf{y}) - \mathbf{g}(\mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|} = \lim_{\mathbf{y} \rightarrow \mathbf{x}} \frac{\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \mathbf{A}(\mathbf{y} - \mathbf{x})}{\|\mathbf{y} - \mathbf{x}\|} = \mathbf{0}. \quad (8.8)$$

Tato limita nejde nijak zjednodušit, jako jsme to udělali s limitou (8.5).

Definice (8.8) není pro ověřování diferencovatelnosti pohodlná, protože počítání limit funkcí více proměnných je obtížné. Uvedeme proto postačující (ne však nutnou) podmínku pro diferencovatelnost, která je mnohem pohodlnější. Zobrazení \mathbf{f} je v bodě \mathbf{x} **spojitě diferencovatelné**, jestliže existují všechny parciální derivace $\partial f_i(\mathbf{x})/\partial x_j$ a funkce $\partial f_i/\partial x_j$ jsou v bodě \mathbf{x} spojité.

Věta 8.3. *Je-li zobrazení v bodě spojité diferencovatelné, je v tomto bodě totálně diferencovatelné.*

Příklad 8.6. Obě parciální derivace (8.6) funkce (8.2) jsou spojité funkce na celém \mathbb{R}^2 (neboť splňují předpoklady Věty 8.1). Proto je funkce (8.2) spojité diferencovatelná na \mathbb{R}^2 , a tedy dle Věty 8.3 je diferencovatelná na \mathbb{R}^2 . ♦

Pozor: pouhá existence všech parciálních derivací pro diferencovatelnost nestačí.

V případě, že je zobrazení \mathbf{f} je v bodě \mathbf{x} diferencovatelné, má matice \mathbf{A} přirozený tvar: její složky jsou parciální derivace všech složek zobrazení podle všech proměnných:

$$\mathbf{A} = \frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} = \mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}. \quad (8.9)$$

Matice (8.9) se nazývá **totální derivace** (nebo krátce jen **derivace**) zobrazení \mathbf{f} v bodě \mathbf{x} . Z historických důvodů se jí říká také **Jacobiho matice**. Funkci, která bodu $\mathbf{x} \in \mathbb{R}^n$ přiřadí (totální) derivaci zobrazení \mathbf{f} v bodě \mathbf{x} , 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:

- Pro $f: \mathbb{R} \rightarrow \mathbb{R}$ je $f'(x)$ skalár a splývá s obyčejnou derivací (8.3) (dle Tvzení 8.2).
- Pro $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^m$ je $\mathbf{f}'(x) = \begin{bmatrix} f'_1(x) \\ \vdots \\ f'_m(x) \end{bmatrix}$ sloupcový vektor, jehož složky jsou derivace složek $f_1, \dots, f_m: \mathbb{R} \rightarrow \mathbb{R}$ zobrazení \mathbf{f} . Funkce je diferencovatelná, právě když jsou všechny její složky diferencovatelné, protože rovnost (8.8) je ekvivalentní m rovnostem (8.5).
- Pro $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je $f'(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \cdots \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right]$ řádkový vektor.

⁴Tato ceta je možná a vede na zajímavou definici totální derivace podle Caratheodoryho.

8.5.1 Derivace složeného zobrazení

Známé ‘řetízkové pravidlo’ pro derivaci složených funkcí jedné proměnné lze přirozeným způsobem rozšířit na zobrazení. Důkaz následující věty není krátký a nebudeme ho uvádět.

Věta 8.4 (řetízkové pravidlo). Necht’ zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je diferencovatelné v bodě $\mathbf{x} \in \mathbb{R}^n$ a zobrazení $\mathbf{g}: \mathbb{R}^m \rightarrow \mathbb{R}^l$ je diferencovatelné v bodě $\mathbf{f}(\mathbf{x})$. Pak složené zobrazení $\mathbf{g} \circ \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^l$ je diferencovatelné v bodě \mathbf{x} a jeho derivace je rovna

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = (\mathbf{g}' \circ \mathbf{f})(\mathbf{x}) \mathbf{f}'(\mathbf{x}) \quad \text{neboli} \quad \frac{d\mathbf{g}(\mathbf{f}(\mathbf{x}))}{d\mathbf{x}} = \mathbf{g}'(\mathbf{f}(\mathbf{x})) \mathbf{f}'(\mathbf{x}). \quad (8.10)$$

Dimenze zúčastněných prostorů lze přehledně znázornit zápisem (viz §1.1.2)

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

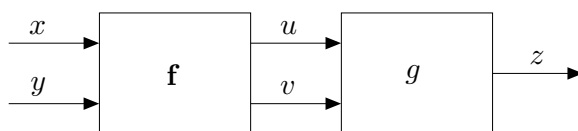
Pokud píšeme $\mathbf{u} = \mathbf{f}(\mathbf{x})$ a $\mathbf{y} = \mathbf{g}(\mathbf{u})$, pravidlo lze psát také v Leibnizově značení jako

$$\frac{d\mathbf{y}}{d\mathbf{x}} = \frac{d\mathbf{y}}{d\mathbf{u}} \frac{d\mathbf{u}}{d\mathbf{x}}, \quad (8.11)$$

což se dobře pamatuje, protože $d\mathbf{u}$ se ‘jakoby vykrátí’ (což ale není důkaz ničeho!). 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 řetízkové pravidlo pro derivaci složení funkcí jedné proměnné. Pravidlo se indukci dá zobecnit na složení více než dvou zobrazení (řetízek): *Jacobiho matice složeného zobrazení je součinem Jacobiho matic jednotlivých zobrazení.*

Příklad 8.7. Necht’ $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je diferencovatelná funkce. Najdi derivaci funkce $h(x, y) = g(x + y, xy)$ podle vektoru (x, y) , tedy řádkový vektor $h'(x, y) = [h_x(x, y) \quad h_y(x, y)] \in \mathbb{R}^{1 \times 2}$.

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 řádkový vektor $g'(u, v) = [g_u(u, v) \quad g_v(u, v)]$. Derivace zobrazení \mathbf{f} podle vektoru (x, y) je matice 2×2

$$\mathbf{f}'(x, y) = \frac{d\mathbf{f}(x, y)}{d(x, y)} = \begin{bmatrix} \frac{\partial(x+y)}{\partial x} & \frac{\partial(x+y)}{\partial y} \\ \frac{\partial(xy)}{\partial x} & \frac{\partial(xy)}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix}.$$

Derivace zobrazení $h = g \circ \mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}$ podle vektoru (x, y) je 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 h_x funkce $h(x, y) = e^{(x+y)^2+(xy)^2}$:

- Považujeme y za konstantu a derivujeme h jako funkci jedné proměnné x :

$$h_x(x, y) = [2(x+y) + 2(xy)y]e^{(x+y)^2+(xy)^2} = 2(x+y+xy^2)e^{(x+y)^2+(xy)^2}.$$

- Položíme $u = x+y$, $v = xy$, $g(u, v) = e^{u^2+v^2}$. Z Příkladu 8.7 máme $h_x = g_u + yg_v$. Jelikož

$$g_u(u, v) = 2ue^{u^2+v^2}, \quad g_v(u, v) = 2ve^{u^2+v^2},$$

$$\text{máme } h_x(x, y) = g_u + yg_v = 2ue^{u^2+v^2} + y(2v)e^{u^2+v^2} = 2(x+y+xy^2)e^{(x+y)^2+(xy)^2}. \quad \blacklozenge$$

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

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

8.5.2 Maticový kalkulus

Je-li nějaké zobrazení zadáno výrazem obsahujícím vektory a matice, jeho derivaci lze obvykle napsat jako výraz obsahující tytéž vektory a matice. To znamená, že při počítání derivací považujeme matice a vektory za *nedělitelné objekty*. Část matematiky zabývající se derivacemi zobrazení s vektory a maticemi je známa jako *maticový (diferenciální) kalkulus*⁵.

Příklad 8.10. Najděme derivaci afinního zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, tedy

$$f_i(\mathbf{x}) = \sum_{j=1}^n a_{ij}x_j + b_i, \quad i = 1, \dots, m.$$

Je $\partial f_i(\mathbf{x})/\partial x_j = a_{ij}$, tedy dle (8.9) je $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$. Tomu se nedivíme, neboť zobrazení \mathbf{f} již je afinní a proto jeho afinní aproximace (8.7) musí být ono samo. \blacklozenge

Příklad 8.11. Najděme derivaci kvadratické formy $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij}x_i x_j, \quad (8.12)$$

kde $\mathbf{A} = [a_{ij}]$ je (ne nutně symetrická) čtvercová matice. Parciální derivace této funkce podle k -té proměnné je (rozmyslete!)

$$\frac{\partial f(\mathbf{x})}{\partial x_k} = \sum_{j=1}^n a_{kj}x_j + \sum_{i=1}^n a_{ik}x_i.$$

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

⁵Asi víte, že *diferenciální počet* (angl. *differential calculus*) se zabývá derivacemi funkcí.

Příklad 8.12. Odvod'eme vektorovou formu známého pravidla pro derivaci součinu. Necht'

$$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x}) = \sum_{i=1}^m g_i(\mathbf{x}) h_i(\mathbf{x}), \quad (8.13)$$

kde $\mathbf{g}, \mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Podle pravidla pro derivaci součinu skalárních funkcí je

$$\frac{\partial f(\mathbf{x})}{\partial x_j} = \sum_{i=1}^m \left(g_i(\mathbf{x}) \frac{\partial h_i(\mathbf{x})}{\partial x_j} + h_i(\mathbf{x}) \frac{\partial g_i(\mathbf{x})}{\partial x_j} \right).$$

Tedy $f'(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$. ◆

Příklad 8.13. Najd'eme derivaci zobrazení $\mathbf{f}(\mathbf{A}\mathbf{x} + \mathbf{b})$ podle \mathbf{x} , kde $\mathbf{f}: \mathbb{R}^l \rightarrow \mathbb{R}^m$ a $\mathbf{A} \in \mathbb{R}^{l \times n}$. V řetízkovém pravidle máme $\mathbb{R}^n \xrightarrow{\mathbf{g}} \mathbb{R}^l \xrightarrow{\mathbf{f}} \mathbb{R}^m$, kde $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$. Je $\mathbf{g}'(\mathbf{x}) = \mathbf{A}$, tedy

$$\frac{d\mathbf{f}(\mathbf{A}\mathbf{x} + \mathbf{b})}{d\mathbf{x}} = \mathbf{f}'(\mathbf{A}\mathbf{x} + \mathbf{b})\mathbf{A}. \quad \blacklozenge$$

Tato a některá další pravidla uvedeme v tabulkách. Nejprve pro zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$:

$\mathbf{f}(\mathbf{x})$	$\mathbf{f}'(\mathbf{x})$	poznámka
\mathbf{x}	\mathbf{I}	
$\mathbf{A}\mathbf{x} + \mathbf{b}$	\mathbf{A}	$\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$
$\mathbf{A}\mathbf{g}(\mathbf{x})$	$\mathbf{A}\mathbf{g}'(\mathbf{x})$	$\mathbf{A} \in \mathbb{R}^{m \times l}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^l$
$\mathbf{g}(\mathbf{A}\mathbf{x} + \mathbf{b})$	$\mathbf{g}'(\mathbf{A}\mathbf{x} + \mathbf{b})\mathbf{A}$	$\mathbf{A} \in \mathbb{R}^{l \times n}$, $\mathbf{g}: \mathbb{R}^l \rightarrow \mathbb{R}^m$

A dále pro některé funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$:

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

Všechny vzorečky v tabulce se vlastně dají odvodit z derivace součinu a řetízkového pravidla (poslední dva řádky). Např. $\mathbf{x}^T \mathbf{A}\mathbf{x} = \mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x})$ kde $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ a $\mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Rozmyslete i pro ostatní řádky tabulky!

Občas se setkáme s funkcemi $\mathbb{R} \rightarrow \mathbb{R}^{m \times n}$, které skalárům přiřazují matice, a s funkcemi $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}$, které maticím přiřazují skaláry (příkladem druhé funkce je stopa, determinant nebo maticová norma). Přísně formálně, naši definici (8.8) totální derivace na taková zobrazení nelze použít, protože je formulována pouze pro zobrazení typu $\mathbb{R}^n \rightarrow \mathbb{R}^m$. Ovšem lze ji použít, když si matici $m \times n$ si představíme 'přerovnanou' do vektoru délky mn .

Složky funkce $\mathbf{F}: \mathbb{R} \rightarrow \mathbb{R}^{m \times n}$ jsou jednoduše funkce $f_{ij}: \mathbb{R} \rightarrow \mathbb{R}$ jediné skalární proměnné. Funkce je diferencovatelná, právě když všechny její složky jsou diferencovatelné. Derivaci této funkce v bodě $x \in \mathbb{R}$ je přirozené definovat jako matici

$$\mathbf{F}'(x) = \begin{bmatrix} f'_{11}(x) & \cdots & f'_{1n}(x) \\ \vdots & \ddots & \vdots \\ f'_{m1}(x) & \cdots & f'_{mn}(x) \end{bmatrix} \in \mathbb{R}^{m \times n}. \quad (8.14)$$

Pro $n = 1$ je tato matice Jacobiho matice (8.9) zobrazení $\mathbb{R} \rightarrow \mathbb{R}^m$, tedy sloupcový vektor.

Derivaci diferencovatelné funkce $f: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ v bodě $\mathbf{X} = [x_{ij}] \in \mathbb{R}^{m \times n}$ je přirozené definovat jako matici

$$f'(\mathbf{X}) = \begin{bmatrix} \frac{\partial f(\mathbf{X})}{\partial x_{11}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial x_{m1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{X})}{\partial x_{1n}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial x_{mn}} \end{bmatrix} \in \mathbb{R}^{n \times m}. \quad (8.15)$$

Matice (8.15) má rozměr $n \times m$ (tedy obráceně, než čekáte), aby pro případ $n = 1$ souhlasila s Jacobiho maticí (8.9). Zde jsou bez důkazů vzorce pro derivace některých funkcí $f: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ (nemusíte se je učit nazpaměť, je to spíše motivace pro další studium maticového počtu):

$f(\mathbf{X})$	$f'(\mathbf{X})$	poznámka
$\text{tr } \mathbf{X} = \langle \mathbf{I}, \mathbf{X} \rangle$	\mathbf{I}	\mathbf{X} čtvercová
$\text{tr}(\mathbf{A}\mathbf{X}) = \langle \mathbf{A}^T, \mathbf{X} \rangle$	\mathbf{A}	
$\ \mathbf{X}\ ^2 = \langle \mathbf{X}, \mathbf{X} \rangle = \text{tr}(\mathbf{X}^T \mathbf{X})$	$2\mathbf{X}^T$	
$\text{tr}(\mathbf{X}^T \mathbf{A}\mathbf{X}) = \langle \mathbf{X}, \mathbf{A}\mathbf{X} \rangle$	$\mathbf{X}^T(\mathbf{A} + \mathbf{A}^T)$	
$\mathbf{a}^T \mathbf{X} \mathbf{b}$	$\mathbf{b} \mathbf{a}^T$	
$\text{tr}(\mathbf{A}\mathbf{X}\mathbf{B})$	$\mathbf{B}\mathbf{A}$	
$\text{tr}(\mathbf{X}^k)$	$k\mathbf{X}^{k-1}$	$k \in \mathbb{N}$
$\text{tr}(\mathbf{X}^{-1})$	$-\mathbf{X}^{-1}\mathbf{X}^{-1} = -\mathbf{X}^{-2}$	\mathbf{X} regulární
$\det \mathbf{X}$	$(\det \mathbf{X})\mathbf{X}^{-1}$	\mathbf{X} regulární
$\ln \det \mathbf{X}$	\mathbf{X}^{-1}	\mathbf{X} pozitivně definitní
$g(\mathbf{X}^T)$	$g'(\mathbf{X}^T)^T$	$g: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$

(*) Derivace zobrazení mezi obecnými lineárními prostory

Definici derivace (8.8) lze zobecnit na zobrazení $f: U \rightarrow V$ mezi obecnými normovanými lineárními prostory U, V : zobrazení $f: U \rightarrow V$ je *diferencovatelné* v bodě $x \in U$, jestliže existuje lineární zobrazení $a: U \rightarrow V$ tak, že

$$\lim_{y \rightarrow x} \frac{\|f(y) - f(x) - a(y - x)\|}{\|y - x\|} = 0, \quad (8.16)$$

kde $\|\cdot\|$ značí ve jmenovateli normu na prostoru U a v čitateli normu na prostoru V . Pak se zobrazení $a = f'(x)$ nazývá totální (nebo Fréchetova) derivace zobrazení f v bodě x . Řetízkové pravidlo zní

$$(g \circ f)'(x) = (g' \circ f)(x) \circ f'(x) = g'(f(x)) \circ f'(x) \quad (8.17)$$

kde $U \xrightarrow{f} V \xrightarrow{g} W$ a $U \xrightarrow{f'(x)} V \xrightarrow{g'(f(x))} W$. Tedy zatímco v (8.10) násobíme Jacobiho matice, v (8.17) skládáme lineární zobrazení.

Každý ze zúčastněných lineárních prostorů U, V, \dots může být prostor skalárů, vektorů nebo matic. Problém je, jak kompaktně reprezentovat lineární zobrazení mezi těmito prostory. Lineární zobrazení $\mathbb{R} \rightarrow \mathbb{R}^m, \mathbb{R}^n \rightarrow \mathbb{R}, \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ a $\mathbb{R} \rightarrow \mathbb{R}^{k \times l}$ lze reprezentovat maticemi: pak násobení matice vektorem lze vidět jako konvenci pro reprezentaci lineárních zobrazení $\mathbb{R}^n \rightarrow \mathbb{R}^m$ a maticový součin jako konvenci pro skládání takových zobrazení. Ovšem lineární zobrazení $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^l, \mathbb{R}^l \rightarrow \mathbb{R}^{m \times n}$ a $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{k \times l}$ nelze reprezentovat maticemi a pro manipulaci s takovými zobrazeními je pohodlnější použít tensorů (jako to dělají fyzikové).

8.6 Směrová derivace

Řez zobrazení $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru⁶ $\mathbf{v} \in \mathbb{R}^n$ je zobrazení $\varphi: \mathbb{R} \rightarrow \mathbb{R}^m$ s hodnotami

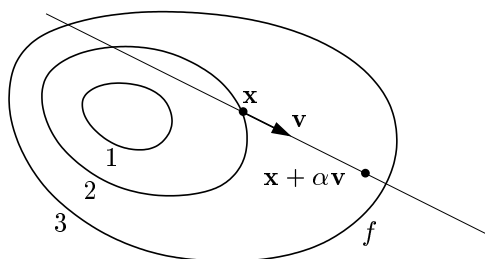
$$\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{v}). \quad (8.18)$$

Směrová derivace⁷ zobrazení f v bodě \mathbf{x} ve směru \mathbf{v} je vektor

$$\mathbf{f}_v(\mathbf{x}) = \varphi'_+(0) = \lim_{\alpha \rightarrow 0^+} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha \rightarrow 0^+} \frac{f(\mathbf{x} + \alpha\mathbf{v}) - f(\mathbf{x})}{\alpha}, \quad (8.19)$$

kde $\varphi'_+(0)$ označuje derivaci zobrazení φ v bodě $\alpha = 0$ zprava, tedy vektor se složkami $\varphi'_{1+}(0), \dots, \varphi'_{m+}(0)$ kde $\varphi'_{i+}(0)$ označuje derivaci funkce φ_i v bodě $\alpha = 0$ zprava. Směrová derivace existuje právě když existují derivace všech těchto složek. Pak také říkáme, že zobrazení f je *diferencovatelné v bodě \mathbf{x} ve směru \mathbf{v}* .

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



Příklad 8.14. Spočítejme směrovou derivaci $f_{(u,v)}(x, y)$ funkce

$$f(x, y) = \sin(x + y^2) \quad (8.20)$$

v bodě (x, y) ve směru (u, v) :

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

Parciální derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ není nic jiného než její směrová derivace ve směru i -tého vektoru standardní báze $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ (jednička na i -tém místě).

Jestliže je zobrazení totálně diferencovatelné, směrová derivace má jednodušší vyjádření:

Věta 8.5. Jestliže je zobrazení $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ totálně diferencovatelné, pak je v tomto bodě diferencovatelné v každém směru $\mathbf{v} \in \mathbb{R}^n$ a platí

$$\mathbf{f}_v(\mathbf{x}) = \mathbf{f}'(\mathbf{x})\mathbf{v}. \quad (8.21)$$

⁶Někdy se řez a směrová derivace definují jen pro *normalizované* vektory \mathbf{v} (tj. $\|\mathbf{v}\| = 1$), protože ty lépe zachycují pojem 'směr'. Většina autorů (a my také) však dovoluje libovolný vektor \mathbf{v} .

⁷Někteří autoři definují směrovou derivaci jako oboustrannou, kde se místo jednostranné limity (8.19) zprava vezme limita oboustranná. Pokud existují derivace zprava $\varphi'_+(0)$ i derivace zleva $\varphi'_-(0)$ a jsou si rovny, pak jednostranná i oboustranná směrová derivace jsou si rovny.

Důkaz. Zobrazení $\mathbf{y} = \varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha\mathbf{v})$ je složením zobrazení $\mathbf{y} = \mathbf{f}(\mathbf{u})$ a $\mathbf{u} = g(\alpha) = \mathbf{x} + \alpha\mathbf{v}$, tedy $\mathbb{R} \xrightarrow{g(\alpha)=\mathbf{u}} \mathbb{R}^n \xrightarrow{\mathbf{f}(\mathbf{u})=\mathbf{y}} \mathbb{R}^m$. Je $g'(\alpha) = \mathbf{v}$. Protože \mathbf{f} má derivaci v bodě \mathbf{x} , podle řetízkového pravidla má také φ (oboustrannou) derivaci v bodě \mathbf{x} a

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

Ale $g(0) = \mathbf{x}$ a proto $\mathbf{f}_{\mathbf{v}}(\mathbf{x}) = \varphi'_+(0) = \varphi'(0) = \mathbf{f}'(\mathbf{x})\mathbf{v}$. ■

Příklad 8.15. Spočítejme směrovou derivaci $f_{(u,v)}(x, y)$ funkce (8.20) podle Věty 8.5:

$$f_{(u,v)}(x, y) = uf_x(x, y) + vf_y(x, y) = u \cos(x + y^2) + 2vy \cos(x + y^2) = (u + 2vy) \cos(x + y^2).$$

To je stejný výsledek jako nám vyšel v Příkladě 8.14. ◆

Rovnost (8.21) říká, že směrová derivace zobrazení \mathbf{f} v pevném bodě \mathbf{x} je *lineární zobrazení* směru \mathbf{v} , reprezentované Jacobiho maticí $\mathbf{f}'(\mathbf{x})$. Speciálně je $\mathbf{f}_{-\mathbf{v}}(\mathbf{x}) = -\mathbf{f}_{\mathbf{v}}(\mathbf{x})$. Zdůrazněme, že nic takového obecně neplatí, pokud zobrazení \mathbf{f} není (totálně) diferencovatelné:

Příklad 8.16. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je $f(\mathbf{x}) = \|\mathbf{x}\|$ (tj. eukleidovská norma). Tato funkce v bodě $\mathbf{x} = \mathbf{0}$ není (totálně) diferencovatelná (pro $n = 1$ je $f(x) = |x|$, nakreslete si graf funkce pro $n = 2$). Spočítejme její směrové derivace v tomto bodě podle definice (8.19). Řez funkce f v bodě $\mathbf{x} = \mathbf{0}$ ve směru $\mathbf{v} \in \mathbb{R}^n$ je

$$\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{v}) = f(\alpha\mathbf{v}) = \|\alpha\mathbf{v}\| = |\alpha| \cdot \|\mathbf{v}\|.$$

Funkce φ v bodě $\alpha = 0$ nemá oboustrannou derivaci, ale má obě jednostranné derivace. Směrová derivace funkce f v bodě $\mathbf{x} = \mathbf{0}$ ve směru \mathbf{v} je $f_{\mathbf{v}}(\mathbf{0}) = \varphi'_+(0) = \|\mathbf{v}\|$, protože derivace funkce $|\alpha|$ v bodě $\alpha = 0$ zprava je 1. Všimněte si, že $\|\mathbf{v}\|$ není lineární funkcí vektoru \mathbf{v} . ◆

8.7 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 vektor⁸.

Zkoumejme směrovou derivaci $f_{\mathbf{v}}(\mathbf{x})$ diferencovatelné funkce f v pevném bodě \mathbf{x} pro různé normalizované směry \mathbf{v} (tedy $\|\mathbf{v}\| = 1$). Dle Věty 8.5 je $f_{\mathbf{v}}(\mathbf{x}) = f'(\mathbf{x})\mathbf{v} = \nabla f(\mathbf{x})^T \mathbf{v}$, což je skalární součin gradientu v bodě \mathbf{x} a vektoru \mathbf{v} . Je jasné (ale promyslete!), že:

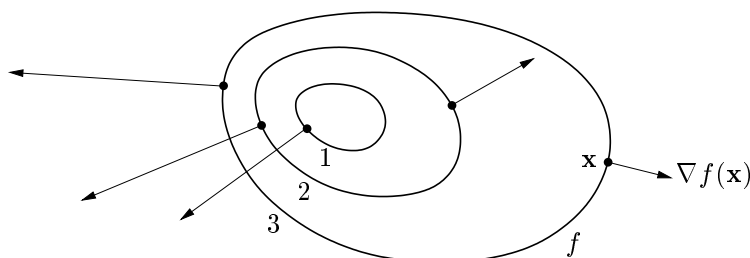
- Směrová derivace je největší ve směru $\mathbf{v} = \nabla f(\mathbf{x}) / \|\nabla f(\mathbf{x})\|$, tedy když je \mathbf{v} rovnoběžný s gradientem a stejně orientovaný. Tedy směr gradientu je *směr největšího růstu* funkce.

⁸Zavedení nového slova pro transpozici derivace se zdá zbytečné – důvod je ale v tom, že totální diferenciál je *lineární forma*, kdežto gradient je *vektor*. Toto rozlišení by bylo důležité, kdybychom uvažovali derivaci funkce $f: V \rightarrow \mathbb{R}$ kde V je obecný (normovaný) lineární prostor (tj. ne nutně \mathbb{R}^n). Různí autoři bohužel nejsou jednotní v rozlišování gradientu a (totální) derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Někdo obojí ztotožňuje a značí $\nabla f(\mathbf{x})$. To ale vede mj. k nekonzistenci s reprezentací totální derivace pomocí Jacobiho matice (8.9), neboť derivace funkce $\mathbb{R}^n \rightarrow \mathbb{R}$ pak není speciálním případem derivace zobrazení $\mathbb{R}^n \rightarrow \mathbb{R}^m$ pro $m = 1$, což je řádkový vektor.

- Velikost gradientu $\|\nabla f(\mathbf{x})\|$ je strmota funkce ve směru největšího růstu.
- Směrová derivace ve směru kolmém na gradient je nulová.

Dále lze ukázat, že za jistých podmínek (viz §11.2.1) je gradient vždy *kolmý k vrstevnici*.

Obrázek ukazuje tři vrstevnice funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a její gradienty v několika bodech:



8.8 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

$$\frac{\partial}{\partial x_j} \frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}.$$

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 8.6. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$. 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.

Příklad 8.17. 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 partiálních 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 8.6 je tato matice symetrická. Často se jí říká **Hessova matice**.

Zde jsou druhé derivace některých funkcí $f: \mathbb{R}^n \rightarrow \mathbb{R}$:

$f(\mathbf{x})$	$f''(\mathbf{x})$	poznámka
$\mathbf{a}^T \mathbf{x} + b$	$\mathbf{0}$	
$\mathbf{x}^T \mathbf{x} = \ \mathbf{x}\ ^2$	$2\mathbf{I}$	
$\mathbf{x}^T \mathbf{A} \mathbf{x}$	$\mathbf{A} + \mathbf{A}^T$	
$g(\mathbf{A} \mathbf{x} + \mathbf{b})$	$\mathbf{A}^T g''(\mathbf{A} \mathbf{x} + \mathbf{b}) \mathbf{A}$	$\mathbf{A} \in \mathbb{R}^{m \times n}, g: \mathbb{R}^m \rightarrow \mathbb{R}$

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 $m \times n \times n$.

8.9 Shrnutí značení derivací

Ted' je správná chvíle shrnout značení všech derivací, které jsme v této kapitole potkali:

symbol	význam
$f'(x), \frac{df(x)}{dx} \in \mathbb{R}$	(oboustranná) derivace funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v bodě $x \in \mathbb{R}$
$f'_+(x), f'_-(x) \in \mathbb{R}$	derivace funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ v bodě $x \in \mathbb{R}$ zprava/zleva
$f_{x_i}(\mathbf{x}), \frac{\partial f(\mathbf{x})}{\partial x_i} \in \mathbb{R}$	partiální derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ v bodě $\mathbf{x} \in \mathbb{R}^n$ podle proměnné x_i
$\mathbf{f}_{\mathbf{v}}(\mathbf{x}) \in \mathbb{R}^m$	směrová derivace zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \mathbb{R}^n$
$\mathbf{f}'(\mathbf{x}), \frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} \in \mathbb{R}^{m \times n}$	(totální) derivace (Jacobiho matice) zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$
$\nabla f(\mathbf{x}) \in \mathbb{R}^n$	gradient funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ v bodě $\mathbf{x} \in \mathbb{R}^n$
$\mathbf{F}'(x), \frac{d\mathbf{F}(x)}{dx} \in \mathbb{R}^{m \times n}$	derivace zobrazení $\mathbf{F}: \mathbb{R} \rightarrow \mathbb{R}^{m \times n}$ v bodě $x \in \mathbb{R}$
$f'(\mathbf{X}), \frac{df(\mathbf{X})}{d\mathbf{X}} \in \mathbb{R}^{n \times m}$	derivace zobrazení $f: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ v bodě $\mathbf{X} \in \mathbb{R}^{m \times n}$
$f''(\mathbf{x}) \in \mathbb{R}^{n \times n}$	druhá derivace (Hessova matice) funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ v bodě $\mathbf{x} \in \mathbb{R}^n$

Podotkněme, že značení derivací nejsou zcela ustálená a jiní lidé mohou používat trošku jiná značení. Zvolili jsme značení, která jsou nejčastěji používaná a dohromady konzistentní.

8.10 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 $T_x^k: \mathbb{R} \rightarrow \mathbb{R}$ takový, že v bodě x má všechny derivace až do řádu k stejné jako funkce f . V tomto smyslu je polynom T_x^k aproximací funkce f v okolí bodu x .

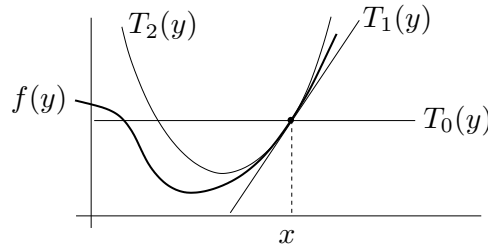
Taylorův polynom je tímto požadavkem definován jednoznačně a má tvar (odvoďte!)

$$T_x^k(y) = \sum_{i=0}^k \frac{1}{i!} f^{(i)}(x) (y - x)^i, \quad (8.22)$$

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 dva:

$$\begin{aligned} T_x^0(y) &= f(x), \\ T_x^1(y) &= f(x) + f'(x)(y-x), \\ T_x^2(y) &= f(x) + f'(x)(y-x) + \frac{1}{2}f''(x)(y-x)^2. \end{aligned}$$

Taylorův polynom nultého stupně T_x^0 je hodně špatná aproximace, rovná jednoduše konstantní funkci. Polynom prvního stupně T_x^1 je vlastně afinní funkce g v (8.4). Polynom druhého stupně T_x^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 $T_{\mathbf{x}}^k: \mathbb{R}^n \rightarrow \mathbb{R}$, který má v bodě \mathbf{x} všechny parciální derivace až do řádu k stejné jako funkce f . Nebudeme uvádět vzorec pro polynom libovolného stupně, budou nám stačit jen polynomy do stupně dva:

$$T_{\mathbf{x}}^0(\mathbf{y}) = f(\mathbf{x}), \quad (8.23a)$$

$$T_{\mathbf{x}}^1(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.23b)$$

$$T_{\mathbf{x}}^2(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})(\mathbf{y} - \mathbf{x}). \quad (8.23c)$$

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 (8.23b) je afinní a funkce (8.23c) 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

$$\mathbf{T}_{\mathbf{x}}^1(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.24)$$

což je vlastně afinní zobrazení \mathbf{g} v (8.7). Polynom druhého stupně vede na zobrazení $\mathbf{T}_{\mathbf{x}}^2$, jehož složky jsou funkce (8.23c). To nejde napsat v maticové formě, protože všech $m \times n \times n$ druhých parciálních derivací se 'nevejde' do matice.

Příklad 8.18. Najděte Taylorův polynom druhého stupně funkce $f(x, y) = x^{-1} + y^{-1} + xy$ v bodě $(x_0, y_0) = (2, 1)$. Máme

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

Dle (8.23c) je (pozor, naše proměnné jsou označeny jinak než v (8.23c))

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

◆

8.11 Cvičení

8.1. Načrtněte několik vrstevnic (přiřpte k nim výšky) těchto funkcí dvou proměnných:

- $f(x_1, x_2) = 2x_1 + x_2$
- $f(x_1, x_2) = x_1 - 3x_2 + 1$
- $f(x_1, x_2) = x_1^2$
- $f(x_1, x_2) = x_1^2 + 4x_2^2$
- $f(x_1, x_2) = x_1^2 - x_2^2 = (x_1 + x_2)(x_1 - x_2)$
- $f(x_1, x_2) = x_1x_2$
- $f(x_1, x_2) = |x_1^2 + x_2^2|$

8.2. 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.

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

- Je funkce f v bodě (x_0, y_0) spojitá?
- Je funkce f v bodě (x_0, y_0) spojitě diferencovatelná?
- Je funkce f v bodě (x_0, y_0) diferencovatelná?
- Najdi totální derivaci (Jacobiho matici) $f'(x, y)$ funkce f v bodě (x_0, y_0) .
- Najdi gradient $\nabla f(x, y)$ funkce f v bodě (x_0, y_0) .
- Najdi řez a směrovou derivaci funkce f v bodě (x_0, y_0) ve směru $(1, -1)$.
- 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 až 17 z Příkladu 8.3.

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

- Spočítej derivaci f podle polárních souřadnic (φ, r) , kde $x = r \cos \varphi$, $y = r \sin \varphi$.
- Bod (x, y) se v čase t pohybuje po křivce dané rovnicí $(x, y) = (t^2 + 2t, \log(t^2 + 1))$. Najděte derivaci f podle času.

8.7. 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 řetízkovým pravidlem.

- a) $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|^2$
- 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) $f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|$
- e) $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$ (kde tedy $f: \mathbb{R}^{2n} \rightarrow \mathbb{R}$ a derivujeme podle vektoru $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2n}$)

8.9. Odvod' totální derivaci kvadratické formy, místo postupu z Příkladu 8.11 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

- a) směr nejstrmějšího stoupání terénu
- b) strmost terénu v jihovýchodním směru.

V této úloze je logické uvažovat směr jako normalizovaný vektor.

8.11. Dokažte, že je-li funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \mathbb{R}^n$, pak pro každé $\alpha \geq 0$ platí $\mathbf{f}_{\alpha \mathbf{v}}(\mathbf{x}) = \alpha \mathbf{f}_{\mathbf{v}}(\mathbf{x})$. Jestliže funkce není totálně diferencovatelná, ukažte, že toto nemusí platit pro $\alpha < 0$.

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

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

8.13. 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.

8.14. Čemu je roven Taylorův polynom prvního stupně polynomu prvního stupně (tj. afinní funkce)? A čemu je roven Taylorův polynom druhého stupně polynomu druhého stupně (tj. kvadratické funkce)? Proč?

8.15. *Metoda konečných diferencí* počítá derivaci funkce přibližně jako

$$f'(x) \approx \frac{f(x + \alpha) - f(x)}{\alpha},$$

kde α 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.

8.16. (*) 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. $\mathbf{f}(\mathbf{x}) = \mathbf{A} \mathbf{x} + \mathbf{x} \mathbf{b}^T \mathbf{x}$). Zamyslete se nad tím, zda lze derivaci $\mathbf{f}'(\mathbf{x})$ vždy vyjádřit výrazem obsahujícím vektor \mathbf{x} , ty samé konstantní matice a vektory, a operace součet matic, maticový součin a transpozice?

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) spojité funkce.

8.3.c) Ano, dle Věty 8.3.

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 $f_{(1, -1)}(x_0, y_0) = 1/3$.

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) $f_t(x, y) = 2f_x(x, y)(t + 1) + 2tf_y(x, y)/(t^2 + 1)$

8.8.e) Nejjednodušší je elementární postup. Výsledek bude řádkový vektor, jehož prvky budou parciální derivace výrazu $\mathbf{x}^T \mathbf{y} = x_1 y_1 + \dots + x_n y_n$ podle $x_1, \dots, x_n, y_1, \dots, y_n$. Tedy $f'(\mathbf{x}, \mathbf{y}) = [\mathbf{y}^T \quad \mathbf{x}^T]$.

Druhá možnost je uvědomit si, že $\mathbf{x}^T \mathbf{y} = \mathbf{z}^T \mathbf{A} \mathbf{z}$ kde $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ a $\mathbf{A} = \frac{1}{2} \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$ a tedy $f'(\mathbf{x}, \mathbf{y}) = f'(\mathbf{z}) = 2\mathbf{z}^T \mathbf{A} = 2[\mathbf{x}^T \quad \mathbf{y}^T] \mathbf{A} = [\mathbf{y}^T \quad \mathbf{x}^T]$.

8.9. Dosad' $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ a $\mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

8.10.a) $(5, 1)/\sqrt{26}$

8.10.b) $(5, 1)^T(1, -1)/\sqrt{2} = 2\sqrt{2}$

8.11. pro $\alpha \geq 0$ dokažte přímo z definice (8.19). Pro $\alpha < 0$ viz Příklad 8.16.

8.12.a) $2e^{-x^2-y^2} \begin{bmatrix} 2x^2 - 1 & 2xy \\ 2xy & 2y^2 - 1 \end{bmatrix}$

8.12.b) $\frac{e^{x+y}}{(e^x+e^y)^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

8.12.c) $f''(\mathbf{x}) = \mathbf{A} + \mathbf{A}^T$

8.12.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)

8.13. $T_{(1, -2)}^0(x, y) = 46$, $T_{(1, -2)}^1(x, y) = 18x - 60y - 92$, $T_{(1, -2)}^2(x, y) = -6x^2 - 24xy - 18x + 24y^2 + 60y + 46$

8.14. V obou případech je roven původnímu polynomu. Ty přece určitě mají nulté a první (a příp. druhé) derivace stejné jako původní polynom.

Kapitola 9

Extrémy funkce na množině

Extrémy funkce na množině jsme už zmínili v §1.2, zde si o nich řekneme více. Kapitola je opakováním některých základních věcí z matematické analýzy a něco navíc.

9.1 Vnitřek a hranice množiny

Označme jako

$$B_\varepsilon(\mathbf{x}) = \{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| < \varepsilon \} \quad (9.1)$$

n -rozměrnou kouli bez hranice¹ se středem $\mathbf{x} \in \mathbb{R}^n$ a poloměrem $\varepsilon > 0$. Všimněte si, že pro speciální případ $n = 1$ je množina (9.1) 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 $B_\varepsilon(\mathbf{x}) \cap (\mathbb{R}^n \setminus X) \neq \emptyset$.

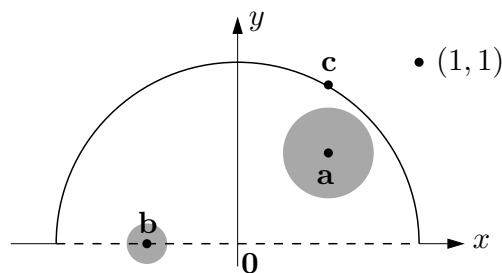
Všimněte si, že hraniční bod množiny nemusí patřit do této množiny. Pokud leží bod v množině, je buď vnitřní nebo hraniční, ale ne obojí najednou (dokažte!). **Vnitřek** [hranice] množiny je množina všech jejích vnitřních [hraničních] bodů.

Množina se nazývá **otevřená**, jestliže všechny její body jsou vnitřní. Množina se nazývá **uzavřená**, jestliže obsahuje každý svůj hraniční bod. Lze dokázat, že množina X je uzavřená [otevřená], právě když její doplněk $\mathbb{R}^n \setminus X$ je otevřený [uzavřený]. Otevřenost a uzavřenost se nevylučují: množiny \emptyset a \mathbb{R}^n jsou zároveň otevřené i uzavřené. Naopak, některé množiny nejsou ani otevřené ani uzavřené, např. interval $(0, 1]$.

Množina X je **omezená**, jestliže existuje $r \in \mathbb{R}$ takové, že $\|\mathbf{x} - \mathbf{y}\| < r$ pro všechna $\mathbf{x}, \mathbf{y} \in X$. Jinými slovy, množina se ‘vejde’ do koule konečného průměru.

Příklad 9.1. Máme množinu $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, y > 0 \} \cup \{(1, 1)\}$ na obrázku:

¹Mohli bychom použít i kouli s hranicí $\{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| \leq \varepsilon \}$. Podobně, norma v (9.1) může být eukleidovská, ale i libovolná jiná vektorová p -norma (viz §12.4.1). Vnitřek a hranice každé množiny by se tím nezměnila.



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.

Množina není otevřená, protože např. bod **c** není vnitřní. Není ani uzavřená, protože např. bod **b** je hraniční ale nepatří do množiny. Množina je omezená. ◆

Příklad 9.2. Bod $1/2$ je vnitřní bod intervalu $(0, 1] \subseteq \mathbb{R}$ a body 0 a 1 jsou hraniční. ◆

Příklad 9.3. Množina $[0, 1] \times \{1\} = \{(x, y) \mid 0 \leq x \leq 1, y = 1\} \subseteq \mathbb{R}^2$ (úsečka v rovině) nemá žádné vnitřní body. Všechny její body jsou hraniční, je tedy sama svou vlastní hranicí. Není otevřená, je uzavřená, je omezená. ◆

Příklad 9.4. 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 $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$. ◆

Příklad 9.5. 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 $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1\}$. ◆

9.2 Existence globálních extrémů

Zopakujme (§1.2), 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 $\mathbf{x}' \in X \setminus \{\mathbf{x}\}$.

Pro odlišení od lokálních minim (viz dále) se tato ‘obyčejná’ minima často nazývají *globální minima*. *Hodnota minima* je pak číslo $f(\mathbf{x}) = \min_{\mathbf{x}' \in X} f(\mathbf{x}')$. Tedy hodnota minima je nejmenší prvek množiny (promyslete!)

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

která je obrazem množiny X v zobrazení f (viz §1.1.2). Funkce na množině nemusí mít minimum, neboť množina $f(X)$ nemusí mít nejmenší prvek.

Pro maxima jsou definice analogické, minima a maxima se dohromady nazývají *extrémy*.

Příklad 9.6. Funkce $f(x) = x$ nemá na množině $(0, 1) \subseteq \mathbb{R}$ (otevřený interval) minimum, neboť množina $f((0, 1)) = (0, 1)$ nemá nejmenší prvek. ◆

Příklad 9.7. Funkce $f(x, y) = x + y$ nemá na množině $X = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ minimum, neboť množina $f(X) = (-\sqrt{2}, \sqrt{2})$ nemá nejmenší prvek. ♦

Uvedeme nyní (bez důkazu) následující zásadní skutečnost:

Věta 9.1. *Spojitě zobrazení uzavřené omezené množiny je uzavřená omezená množina.*

Tedy máme-li uzavřenou a omezenou množinu $X \subseteq \mathbb{R}^n$ a spojitě zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, pak množina $\mathbf{f}(X) = \{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in X\} \subseteq \mathbb{R}^m$ bude také uzavřená a omezená².

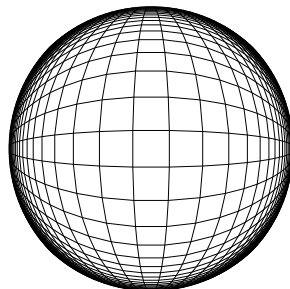
Mohlo by se zdát, že spojitě zobrazení bude zachovávat uzavřenost bez omezenosti nebo omezenost bez uzavřenosti. Je snadné najít protipříklady.

Příklad 9.8. Necht' X je interval $[1, +\infty) \subseteq \mathbb{R}$. Tato množina je uzavřená a není omezená. Spojitě zobrazení $f(x) = 1/x$ zobrazí tuto množinu na množinu $f(X) = (0, 1]$, která není uzavřená a je omezená. ♦

Příklad 9.9. Uvažujme spojitě zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ s hodnotami

$$\mathbf{f}(\mathbf{x}) = \frac{\mathbf{x}}{\sqrt{1 + \mathbf{x}^T \mathbf{x}}}. \quad (9.2)$$

Množina \mathbb{R}^n je uzavřená, otevřená a neomezená. Množina $\mathbf{f}(\mathbb{R}^n)$ je otevřená a omezená: je to jednotková koule (bez hranice). Pro ilustraci je na obrázku množina $\mathbf{f}(X)$ pro $X = (\mathbb{R} \times \mathbb{Z}) \cup (\mathbb{Z} \times \mathbb{R}) \subseteq \mathbb{R}^2$ (tedy X je mřížka v rovině, rozmyslete!):



Věta 9.1 má důležitý důsledek pro optimalizaci, který je znám jako *věta o extrémní hodnotě* nebo *Weierstrassova věta*.

Důsledek 9.2 (Weierstrassova věta). *Spojitá funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nabývá na uzavřené omezené množině $X \subseteq \mathbb{R}^n$ svého minima.*

Důkaz. Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je obraz uzavřené omezené množiny $X \subseteq \mathbb{R}^n$ uzavřená omezená množina $f(X) \subseteq \mathbb{R}$. To ale nemůže být nic jiného než uzavřený konečný interval nebo sjednocení takových intervalů. Taková množina jistě má nejmenší prvek. ■

Zdůrazněme, že Weierstrassova věta udává pouze postačující (avšak ne nutné) podmínky pro existenci minima funkce na množině. Např. funkce $f(x) = x^2$ má na množině \mathbb{R} minimum, i když množina \mathbb{R} není omezená.

²Množinám, které jsou zároveň uzavřené a omezené, se říká také *kompaktní*.

9.3 Lokální extrémy

Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nabývá na množině $X \subseteq \mathbb{R}^n$ v bodě $\mathbf{x} \in X$ svého **lokálního minima**, jestliže existuje $\varepsilon > 0$ tak, že funkce f nabývá na množině $X \cap B_\varepsilon(\mathbf{x})$ svého (globálního) minima (tuto definici promyslete, je důležitá!). Analogicky můžeme definovat ostrá lokální 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.

(Můžeme přirozeně mluvit i o lokálních minimech/maximem/optimech optimalizační úlohy ve standardním tvaru (1.10), což jsou lokální minima/maxima účelové funkce f úlohy na množině (1.12) jejích přípustných řešení.)

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 se nazývá **volný** když \mathbf{x} je vnitřní bod množiny X , a **vázaný** (množinou X) když \mathbf{x} je hraniční bod množiny X .

Tvrzení 9.3. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a necht' \mathbf{x} je vnitřní bod množiny $X \subseteq \mathbb{R}^n$.

Následující výroky jsou ekvivalentní:

- \mathbf{x} je lokální minimum [maximum] funkce f na množině X .
- \mathbf{x} je lokální minimum [maximum] funkce f na množině \mathbb{R}^n .

Důkaz. Necht' $\mathbf{x} \in X$ je lokální minimum f na \mathbb{R}^n , tedy (globální) minimum f na $B_\varepsilon(\mathbf{x})$. Pak ovšem je \mathbf{x} také (globální) minimum f na množině $X \cap B_\varepsilon(\mathbf{x})$, tedy lokální minimum f na X .

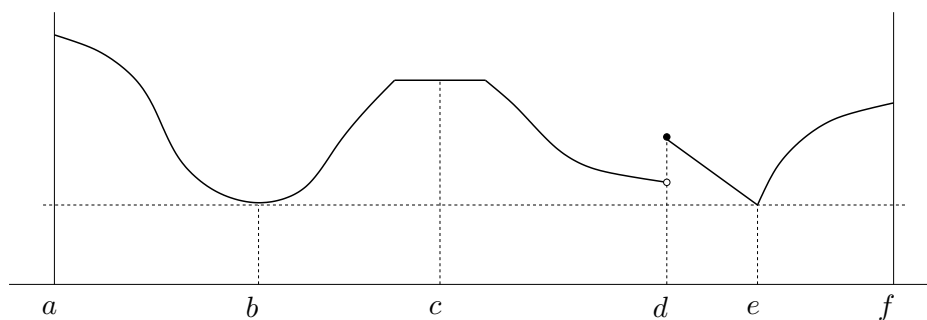
Necht' \mathbf{x} je lokální minimum f na X , tedy (globální) minimum f na množině $X \cap B_\varepsilon(\mathbf{x})$. Protože \mathbf{x} je vnitřní bod množiny X , můžeme ε zmenšit tak, že $B_\varepsilon(\mathbf{x}) \subseteq X$. Pak \mathbf{x} bude stále (globální) minimum f na množině $X \cap B_\varepsilon(\mathbf{x}) = B_\varepsilon(\mathbf{x})$ a tedy lokální minimum f na \mathbb{R}^n .

Důkaz zjevně zůstane v platnosti, nahradíme-li minima maximy. ■

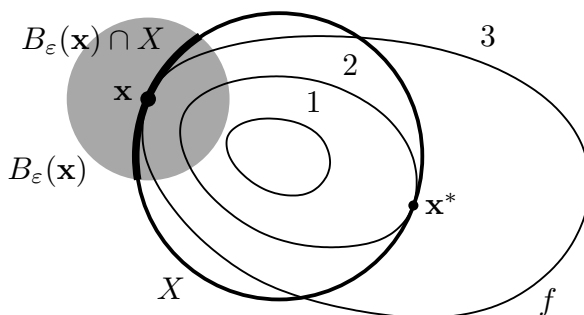
Tvrzení 9.3 ukazuje, že není velký rozdíl mezi volnými lokálními extrémy na množině X a lokálními extrémy na \mathbb{R}^n . Říká, že bod je volný lokální extrém f na X , právě když je to lokální extrém f na \mathbb{R}^n a zároveň vnitřní bod množiny X . Tedy abychom našli volné lokální extrémy f na X , stačí najít lokální extrémy f a zahodit ty, které nejsou vnitřními body množiny X .

Může nastat speciální situace, kdy bod je lokální minimum [maximum] f na \mathbb{R}^n a zároveň hraniční bod množiny X . Pak, dle definice výše, považujeme tento bod za vázané lokální minimum [maximum] f na X , i když ho množina X vlastně nijak neomezuje.

Příklad 9.10. Funkce jedné proměnné na obrázku nabývá na uzavřeném intervalu $[a, f] \subseteq \mathbb{R}$ v bodě a globálního (a tedy i lokálního) maxima, v bodech b, e globálního (a tedy i lokálního) minima, v bodě c lokálního maxima a zároveň lokálního minima, v bodě d lokálního maxima, 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 9.11. 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 extrémů jsou vázané, protože body \mathbf{x}^* a \mathbf{x} jsou hraniční body množiny X (množina X vnitřní body nemá, viz Příklad 9.4). ♦

Příklad 9.12. Funkce $f(\mathbf{x}) = x_1$ na množině $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1\}$ má v bodě $(-1, 0, \dots, 0)$ vázané globální (a tedy i lokální) minimum. ♦

Příklad 9.13. 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. ♦

9.4 Cvičení

9.1. Máme množiny $X = [-1, 1] \times \{0\}$ a $Y = [-1, 1] \times [-1, 1]$. Načrtněte následující množiny:

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

9.2. Načrtněte následující množiny (jedná se o podmnožiny \mathbb{R}^2 nebo \mathbb{R}^3):

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

9.3. Co je vnitřek a hranice těchto množin?

- Množina reálných čísel \mathbb{R}
- Uzavřený interval $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$
- Množina racionálních čísel \mathbb{Q}
- Množina (9.1)

- 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)

9.4. 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) $\{ x + y \mid |x| \geq 1, |y| \geq 1 \}$

9.5. (★) Je-li \mathbf{f} zobrazení z Příkladu 9.9, řekli jsme, že množina $\mathbf{f}(\mathbb{R}^n) = \{ \mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n \}$ obrazů tohoto zobrazení je otevřená jednotková koule, kterou (dle (9.1)) označme jako $B_1(\mathbf{0}_n)$. Dokažte (viz §1.1.2), že

1. Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow B_1(\mathbf{0}_n)$ s hodnotami danými vzorečkem (9.2) je bijektivní.
2. Najděte zobrazení k němu inverzní.

9.6. Najděte všechny (globální i lokální) extrémy funkce $f(x, y) = x + y$ na množině $\{ (x, y) \in \mathbb{R}^2 \mid y = x^2, y \leq 1 \}$.

9.7. 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.

9.8. 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.

- 9.9. Může se stát, že funkce má na množině lokální minimum ale nemá na ní globální minimum? Odpověď dokažte.
- 9.10. Může nekonstantní lineární funkce nabývat na množině lokálního extrému ve vnitřním bodě této množiny? Odpověď dokažte.
- 9.11. Dokažte, že pro libovolné funkce $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ platí

$$\min_{\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n} [f_1(\mathbf{x}_1) + \dots + f_m(\mathbf{x}_m)] = \min_{\mathbf{x} \in \mathbb{R}^n} f_1(\mathbf{x}) + \dots + \min_{\mathbf{x} \in \mathbb{R}^n} f_m(\mathbf{x}), \quad (9.3)$$

tedy minimalizace funkce $f(\mathbf{x}_1, \dots, \mathbf{x}_m) = f_1(\mathbf{x}_1) + \dots + f_m(\mathbf{x}_m)$ se rozpadne na m nezávislých minimalizací.

(\star) Platí tvrzení i tehdy, když nahradíme součet maximem nebo minimem? Obecněji, pro jaké funkce $g: \mathbb{R}^m \rightarrow \mathbb{R}$ platí

$$\min_{\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n} g(f_1(\mathbf{x}_1), \dots, f_m(\mathbf{x}_m)) = g(\min_{\mathbf{x} \in \mathbb{R}^n} f_1(\mathbf{x}), \dots, \min_{\mathbf{x} \in \mathbb{R}^n} f_m(\mathbf{x}))? \quad (9.4)$$

- 9.12. Platilo by Tvrzení 9.3, kdyby se v něm slovo ‘lokální’ nahradilo slovem ‘globální’? Odpověď dokažte.

Nápověda a řešení

9.3.c) Vnitřek je \emptyset a hranice je \mathbb{R} .

9.5. Inverze k \mathbf{f} je zobrazení, tj. $\mathbf{f}^{-1}: B_1(\mathbf{0}_n) \rightarrow \mathbb{R}^n$ s hodnotami $\mathbf{f}^{-1}(\mathbf{y}) = \frac{\mathbf{y}}{\sqrt{1 - \mathbf{y}^T \mathbf{y}}}$. To dokážeme ověřením, že $\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$ pro $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y}$ pro $\mathbf{y} \in B_1(\mathbf{0}_n)$, což uděláme dosazením vzorečků do sebe. Bijektivnost obou zobrazení plyne z existence inverze.

9.6. $(1, 1)$ je glob. maximum, $(-\frac{1}{2}, \frac{1}{4})$ glob. minimum, $(-1, 1)$ lok. maximum. To je vidět z obrázku (nakreslete!). Jiný postup je eliminovat proměnnou y , čímž úlohu převedeme na hledání extrémů funkce $x + x^2$ na intervalu $[-1, 1]$.

9.8. Druhý výrok plyne z prvního. Naopak to ale neplatí, protipříklad je $X = \mathbb{R}$, $X' = [0, 1]$, $f(x) = x$.

Kapitola 10

Volné lokální extrémý

Zde se budeme věnovat podmínkám a algoritmům na volné lokální extrémý diferencovatelných funkcí. Díky Tvzení 9.3 místo volných lokálních extrémů funkcí $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině $X \subseteq \mathbb{R}^n$ budeme uvažovat pouze lokální extrémý funkce f na množině \mathbb{R}^n .

10.1 Analytické podmínky

Je-li směrová derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ v bodě \mathbf{x} ve směru \mathbf{v} kladná, tj. $f_{\mathbf{v}}(\mathbf{x}) > 0$, funkce v bodě \mathbf{x} ve směru \mathbf{v} stoupá. Podobně, je-li $f_{\mathbf{v}}(\mathbf{x}) < 0$, funkce v bodě \mathbf{x} ve směru \mathbf{v} klesá. Upřesněme, co to znamená:

Tvzení 10.1. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \mathbb{R}^n$.

- Jestliže $f_{\mathbf{v}}(\mathbf{x}) > 0$, pak existuje $\delta > 0$ tak, že pro každé $\alpha \in (0, \delta]$ je $f(\mathbf{x} + \alpha\mathbf{v}) > f(\mathbf{x})$.
- Jestliže $f_{\mathbf{v}}(\mathbf{x}) < 0$, pak existuje $\delta > 0$ tak, že pro každé $\alpha \in (0, \delta]$ je $f(\mathbf{x} + \alpha\mathbf{v}) < f(\mathbf{x})$.

Důkaz. Zopakujme definici směrové derivace (8.19):

$$f_{\mathbf{v}}(\mathbf{x}) = \lim_{\alpha \rightarrow 0^+} g(\alpha) = a, \quad \text{kde} \quad g(\alpha) = \frac{f(\mathbf{x} + \alpha\mathbf{v}) - f(\mathbf{x})}{\alpha}.$$

Dle definice limity zprava (viz každá učebnice analýzy) pro každé $\varepsilon > 0$ existuje $\delta > 0$ tak, že pro každé $\alpha \in (0, \delta]$ platí $|g(\alpha) - a| \leq \varepsilon$. Jestliže $a > 0$, existuje tedy $\delta > 0$ tak, že pro každé $\alpha \in (0, \delta]$ platí $g(\alpha) > 0$, tedy $f(\mathbf{x} + \alpha\mathbf{v}) - f(\mathbf{x}) > 0$. Příklad $a < 0$ se dokáže obdobně. ■

Jestliže $f_{\mathbf{v}}(\mathbf{x}) > 0$ [$f_{\mathbf{v}}(\mathbf{x}) < 0$], směr \mathbf{v} nazveme **vzestupný** [sestupný] **směr** (angl. *ascent* [*descent*] *direction*) funkce f v bodě \mathbf{x} . Dle Tvzení 10.1 můžeme v tom případě hodnotu funkce zvětšit [zmenšit] malým posunutím bodu \mathbf{x} ve směru \mathbf{v} . Z toho ihned plyne nutná podmínka na volné lokální extrémý:

Věta 10.2. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \mathbb{R}^n$.

- Jestliže \mathbf{x} je lokální minimum funkce f , pak $f_{\mathbf{v}}(\mathbf{x}) \geq 0$.
- Jestliže \mathbf{x} je lokální maximum funkce f , pak $f_{\mathbf{v}}(\mathbf{x}) \leq 0$.

Důkaz. Jestliže $f_{\mathbf{v}}(\mathbf{x}) < 0$, pak dle Tvzení 10.1 pro každé $\varepsilon > 0$ existuje $\alpha > 0$ tak, že $\mathbf{x} + \alpha\mathbf{v} \in B_{\varepsilon}(\mathbf{x})$ a $f(\mathbf{x} + \alpha\mathbf{v}) < f(\mathbf{x})$, tedy \mathbf{x} není lokální minimum funkce f . Obdobně pro maximum. ■

Větu 10.2 moc často nepoužijete, protože málokdy potkáte funkci, která by v bodě byla diferencovatelná v nějakém směru ale nebyla v něm (totálně) diferencovatelná. Je-li funkce (totálně) diferencovatelná, podmínka na volné lokální extrémy se zjednoduší:

Věta 10.3. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je v bodě $\mathbf{x} \in \mathbb{R}^n$ (totálně) diferencovatelná. Jestliže \mathbf{x} je lokální extrém (tj. minimum či maximum) funkce f , pak $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T = \mathbf{0}$.

Důkaz. Protože f je v bodě \mathbf{x} diferencovatelná, dle Věty 8.5 je $f_{\mathbf{v}}(\mathbf{x}) = f'(\mathbf{x})\mathbf{v}$. Jestliže $f'(\mathbf{x}) \neq \mathbf{0}$, pak existuje směr $\mathbf{v} \in \mathbb{R}^n$ tak, že $f_{\mathbf{v}}(\mathbf{x}) = f'(\mathbf{x})\mathbf{v} > 0$ a $f_{-\mathbf{v}}(\mathbf{x}) = -f'(\mathbf{x})\mathbf{v} < 0$. Dle Věty 10.2 tedy \mathbf{x} není lokální minimum ani maximum. ■

Jestliže je funkce f v bodě \mathbf{x} totálně diferencovatelná a platí $f'(\mathbf{x}) = \mathbf{0}$ (tj. všechny parciální derivace funkce f v bodě \mathbf{x} jsou nulové), bod \mathbf{x} se nazývá **stacionární bod** funkce f . Věta 10.3 říká, že stacionární body jsou body 'podezřelé' z volného lokálního extrému. Věta ovšem svádí k použití v situacích, kdy nejsou splněny její předpoklady. Uved'me příklady těchto situací:

Příklad 10.1. Funkce $f(x) = x^3$ má v bodě 0 stacionární bod, ale nemá tam lokální extrém. ♦

Příklad 10.2. V Příkladu 9.10 jsou předpoklady Věty 10.3 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. ♦

Příklad 10.3. 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 totálně diferencovatelná, a tedy nelze použít Větu 10.3 (lze ovšem použít Tvrzení 10.1, protože f je v bodě $\mathbf{0}$ diferencovatelná ve všech směrech). 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 . ♦

Věta 10.3 udává podmínku *prvního řádu* na volné lokální extrémy, protože obsahuje první derivace. Následující podmínka *druhého řádu* pomůže zjistit, zda je stacionární bod volným lokálním extrémem, případně jakým:

Věta 10.4. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je v bodě $\mathbf{x} \in \mathbb{R}^n$ dvakrát diferencovatelná.

- Jestliže \mathbf{x} je lokální minimum [maximum] funkce f , pak platí $f'(\mathbf{x}) = \mathbf{0}$ a Hessova matice $f''(\mathbf{x})$ je pozitivně [negativně] semidefinitní.
- Jestliže platí $f'(\mathbf{x}) = \mathbf{0}$ a Hessova matice $f''(\mathbf{x})$ je pozitivně [negativně] definitní, pak \mathbf{x} je ostré lokální minimum [maximum] funkce f .

Všimněte si:

- Z Věty 10.4 plyne, že když $f''(\mathbf{x})$ je indefinitní (tedy není ani pozitivně semidefinitní ani negativně semidefinitní), pak \mathbf{x} není lokální extrém. Bod \mathbf{x} , ve kterém $f'(\mathbf{x}) = \mathbf{0}$ a matice $f''(\mathbf{x})$ je indefinitní, se nazývá **sedlový bod** funkce f .
- 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. Příklady jsou funkce x^3 a x^4 v bodě 0.

Větu 10.4 nebudeme dokazovat, uvedeme jen důvod, díky kterému jí snad ochotněji uvěříte. Místo funkce f vyšetřujeme v blízkosti bodu \mathbf{x} její Taylorův polynom druhého stupně (8.23c),

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

Protože $f'(\mathbf{x}) = \mathbf{0}$, lineární člen je nulový a polynom je tedy kvadratická forma posunutá do bodu \mathbf{x} . Zda funkce $T_{\mathbf{x}}^2$ má či nemá v bodě \mathbf{x} extrém bychom tedy mohli určit podle Věty 6.7 z definitnosti matice $f''(\mathbf{x})$. Zde ovšem vyšetřujeme funkci f a ne její aproximaci $T_{\mathbf{x}}^2$, proto pro lokální extrém nestačí (positivní či negativní) semidefinitnost $f''(\mathbf{x})$.

Příklad 10.4. Extrémy kvadratické funkce (6.15) umíme hledat pomocí rozkladu na čtverec. Ovšem je to také možné pomocí derivací. Podmínka stacionarity je

$$\frac{d}{d\mathbf{x}}(\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c) = 2\mathbf{x}^T \mathbf{A} + \mathbf{b}^T = \mathbf{0}.$$

Po transpozici dostaneme rovnici (6.17a). 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. ♦

Podotkněme, že ověřování podmínek druhého řádu pro funkce mnoha proměnných může být nepříjemné a často se snažíme tomu vyhnout (už nalézt Hessián může být problém, natož ověřovat jeho definitnost). Místo toho se snažíme vymyslet jiný (jednodušší) důkaz, že daný stacionární bod je/není lokální minimum/maximum.

10.2 Iterační metody na volné lokální extrémy

Dále se budeme věnovat numerickým iteračním metodám¹ na hledání volných lokálních extrémů (budeme uvažovat pouze minima) diferencovatelných funkcí $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Počínaje počátečním odhadem $\mathbf{x}_0 \in \mathbb{R}^n$ řešení, iterační metoda postupně vytváří posloupnost $\mathbf{x}_1, \mathbf{x}_2, \dots$, která za příznivých okolností konverguje k řešení úlohy (tj. k lokálnímu minimu).

Bod \mathbf{x}_{k+1} závisí na předchozím bodu \mathbf{x}_k a hodnotě účelové funkce f a příp. (pokud je funkce diferencovatelná) na jejích derivacích v tomto bodě, někdy též na hodnotě příp. derivacích v několika minulých bodech $\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots$. **Řádem metody** se myslí nejvyšší řád použité derivace: metody *nultého řádu* (také zvané *metody bez derivací*, *derivative-free methods*) používají jen funkční hodnoty, metody *prvního řádu* navíc první derivace, a metody *druhého řádu* navíc druhé derivace. Metody vyššího než druhého řádu se užívají zřídka.

Základní otázka je, zda metoda konverguje k (nějakému) lokálnímu minimu a když ano, tak jak rychle. Tímto se zabývá *konvergenční analýza* metody a příslušné matematické věty jsou obvykle složité a jejich důkazy ještě složitější. Konvergenční vlastnosti konkrétní metody zjevně závisí na vlastnostech funkce f a na volbě počátečního odhadu \mathbf{x}_0 . Metody vyššího řádu obvykle konvergují k lokálnímu minimu pro menší množinu funkcí f a počátečních odhadů \mathbf{x}_0 , ale když konvergují, tak konvergují obvykle rychleji než metody nižšího řádu.

Zde se zaměříme na třídu metod zvaných **sestupné metody**², jejichž iterace má tvar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{v}_k, \tag{10.1}$$

¹Schválně píšeme *metody* a ne *algoritmy*, neboť algoritmus by měl skončit po konečném počtu operací, kdežto iterační metoda typicky pouze konverguje v nekonečném počtu iterací.

²Existují i metody, ve kterých směr \mathbf{v}_k není vždy sestupný (např. *subgradientní metody*).

kde $\mathbf{v}_k \in \mathbb{R}^n$ je sestupný směr v bodě \mathbf{x}_k (tedy $f_{\mathbf{v}_k}(\mathbf{x}_k) < 0$) zvaný **směr hledání** a skalár $\alpha_k > 0$ je **délka kroku** v k -té iteraci. Metoda v každé iteraci zvolí směr hledání a délku kroku a provede aktualizaci odhadu dle (10.1). Volba směru hledání je klíčový znak pro rozlišení jednotlivých metod.

10.2.1 Volba délky kroku

Máme-li směr hledání \mathbf{v}_k , zde jsou nejběžnější způsoby jak volit délku kroku α_k :

- *Optimální délka kroku.* Je-li směr \mathbf{v}_k sestupný, dle Tvzení 10.1 existuje délka kroku $\alpha_k > 0$ tak, že $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$. Nejlepší délka kroku α_k je minimum řezové funkce

$$\varphi(\alpha) = f(\mathbf{x}_k + \alpha \mathbf{v}_k) \quad (10.2)$$

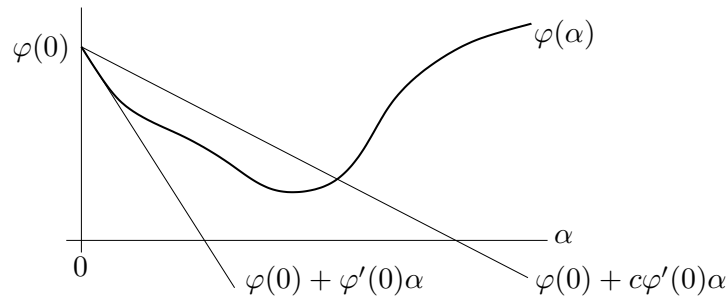
na intervalu $(0, \infty)$. Tato úloha je v kontextu vícerozměrné optimalizace nazývána *jedno-rozměrné hledání* (angl. *line search*).

- *Přibližně optimální délka kroku.* Obvykle není chytré hledat globální minimum (na intervalu $(, \infty)$) funkce (10.2) (leďaže by její řešení šlo hledat velmi levně), protože se stejně v příští iteraci pohneme jinam. Proto se často řeší přibližně. Na nějaký algoritmus hledající přibližné minimum funkce jedné proměnné byste snadno přišli. Ovšem ne každý takový algoritmus garantuje dobré konvergenční vlastnosti iterační metody.

Oblíbená metoda zajišťující dobré konvergenční vlastnosti je *backtracking line search*. Hledáme α splňující tzv. *Armijo-Goldsteinovo pravidlo*: pro nějakou konstantu $c \in (0, 1)$ musí být

$$f(\mathbf{x}_k + \alpha \mathbf{v}_k) = \varphi(\alpha) \leq f(\mathbf{x}_k) + c\alpha f'(\mathbf{x}_k) \mathbf{v}_k = \varphi(0) + c\alpha \varphi'(0). \quad (10.3)$$

Všimněte si, že výraz $\varphi(0) + \alpha \varphi'(0)$ je Taylorův polynom prvního stupně funkce φ v bodě 0 (tedy afinní funkce proměnné α), tedy $\varphi(0) + c\alpha \varphi'(0)$ je tato afinní funkce s mírnějším směrnici, viz obrázek:



Toto pravidlo samo nestačí, protože každá dostatečně malá hodnota $\alpha > 0$ ho splňuje. Proto se užívá v jednoduchém algoritmu: na začátku zvolíme nějakou maximální hodnotu α a parametry $c, \tau \in (0, 1)$ a pak zmenšujeme α jeho násobením číslem τ , dokud nezačne platit (10.3).

- *Pevná posloupnost délek kroku.* V tomto případě délky kroku $\alpha_k > 0$ závisejí pouze na k . Obvykle se vyžaduje, aby posloupnost (α_k) splňovala

$$\lim_{k \rightarrow \infty} \alpha_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty. \quad (10.4)$$

Druhá podmínka je nutná, protože vzdálenost $\|\mathbf{x}_0 - \mathbf{x}^*\|$ mezi počátečním odhadem a hledaným řešením může být libovolně velká. Všimněte si, že i když jsou směry \mathbf{v}_k sestupné, pevná posloupnost délek kroku nezajišťuje monotónní pokles funkce f (tedy pro nějaká k může být $f(\mathbf{x}_{k+1}) > f(\mathbf{x}_k)$).

- *Konstantní délka kroku.* Délka kroku je konstantní, $\alpha_k = \alpha > 0$ pro každé k .

10.3 Gradientní metoda

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

$$\mathbf{v}_k = -f'(\mathbf{x}_k)^T = -\nabla f(\mathbf{x}_k). \quad (10.5)$$

Tento směr je sestupný, což je okamžitě vidět dosazením do (8.21).

Výhodou gradientní metody je spolehlivost, daná tím, že směr (10.5) je vždy sestupný. Nevýhodou je často pomalá konvergence. Lze dokázat, že za jistých nepříliš omezujících předpokladů gradientní metoda v okolí lokálního minima konverguje tzv. *lineárně* (tj. vzdálenost od lokálního minima klesá geometrickou řadou). Ovšem lineární konvergence může být někdy pomalá: to 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).

Příklad 10.5. 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 (10.2) je k -tá iterace rovna (odvoďte!)

$$x_k = \left(-\frac{a-1}{a+1}\right)^k, \quad y_k = a \left(\frac{a-1}{a+1}\right)^k. \quad (10.6)$$

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

10.3.1 Závislost na lineární transformaci souřadnic

Transformujme vektor proměnných \mathbf{x} lineární transformací $\mathbf{y} = \mathbf{A}\mathbf{x}$, kde \mathbf{A} je regulární matice. Je jasné, že funkce f původních proměnných \mathbf{x} bude mít stejné extrémy jako funkce

$$g(\mathbf{y}) = f(\mathbf{x}) = f(\mathbf{A}^{-1}\mathbf{y}).$$

Iterace gradientní metody v nových proměnných je

$$\mathbf{y}_{k+1} = \mathbf{y}_k - \alpha_k g'(\mathbf{y}_k)^T. \quad (10.7)$$

Zjistíme, jaké iteraci to odpovídá v původních proměnných. Z řetízkového pravidla máme

$$g'(\mathbf{y}) = f'(\mathbf{A}^{-1}\mathbf{y})\mathbf{A}^{-1} = f'(\mathbf{x})\mathbf{A}^{-1}.$$

Dosazením za \mathbf{y} a $g'(\mathbf{y})$ do (10.7) a úpravou (proved'te!) dostaneme

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k (\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (10.8)$$

To lze napsat ve tvaru (10.1) se směrem hledání

$$\mathbf{v}_k = -(\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (10.9)$$

Tento směr se liší od původního směru (10.5) vynásobením maticí $(\mathbf{A}^T \mathbf{A})^{-1}$. Vidíme, že gradientní metoda *není invariantní* vůči lineární transformaci souřadnic.

Nový směr (10.9) je také sestupný, tj. $-f'(\mathbf{x}_k)(\mathbf{A}^T \mathbf{A})^{-1}f'(\mathbf{x}_k)^T < 0$, neboť matice $\mathbf{A}^T \mathbf{A}$ a tedy i její inverze je pozitivně definitní (viz Cvičení 6.20 a 6.19).

Na vzorec (10.9) 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 metody uvedené dále budou mít vždy tento tvar sestupného směru, ovšem matice \mathbf{C}_k bude jiná v každém kroku.

10.4 Newtonova metoda

Newtonova metoda (přesněji Newton-Raphsonova, také se jí říká *metoda tečen*) je slavná iterační metoda na řešení soustav nelineárních rovnic. Lze ho použít i na minimalizaci funkce tak, že hledáme její bod s nulovým gradientem. Oba způsoby použití popíšeme.

10.4.1 Použití na soustavy nelineárních rovnic

Necht' $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ je diferencovatelné zobrazení. Chceme řešit rovnici $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, což je soustava n rovnic s n neznámými. Myšlenka Newtonovy metody je jednoduchá: místo hledání nulového bodu zobrazení \mathbf{g} (což je obecně velmi obtížné) opakujeme iteraci, která najde nulový bod afinní aproximace zobrazení \mathbf{g} v okolí aktuálního odhadu (což je snadné).

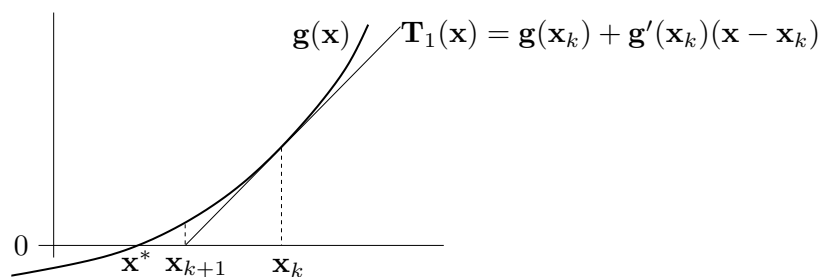
Afinní aproximace zobrazení \mathbf{g} v okolí bodu \mathbf{x}_k je zobrazení (viz (8.7) nebo (8.24))

$$\mathbf{T}_{\mathbf{x}_k}^1(\mathbf{x}) = \mathbf{g}(\mathbf{x}_k) + \mathbf{g}'(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k). \quad (10.10)$$

Další odhad \mathbf{x}_{k+1} najdeme řešením nehomogenní lineární soustavy $\mathbf{T}_{\mathbf{x}_k}^1(\mathbf{x}_{k+1}) = \mathbf{0}$. Pokud je Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ regulární, řešením je

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^{-1}\mathbf{g}(\mathbf{x}_k). \quad (10.11)$$

Viz obrázek:



Hlavní výhodou Newtonovy metody je, že v blízkém okolí řešení obvykle konverguje velmi rychle, tzv. *superlineárně*, mnohem rychleji než gradientní metoda). Nevýhodou je, že je obvykle nutno začít s poměrně přesnou aproximací \mathbf{x}_0 skutečného řešení, jinak metoda snadno diverguje.

Příklad 10.6. *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,

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)} = x_k - \frac{x_k^2 - a}{2x_k} = x_k - \frac{1}{2} \left(x_k - \frac{a}{x_k} \right) = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right). \quad \blacklozenge$$

Příklad 10.7. Hledejme 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

$$\mathbf{g}(x, y) = \begin{bmatrix} (x-1)^2 + y^2 - 1 \\ x^4 + y^4 - 1 \end{bmatrix} \in \mathbb{R}^2, \quad \mathbf{g}'(x, y) = \begin{bmatrix} 2(x-1) & 2y \\ 4x^3 & 4y^3 \end{bmatrix} \in \mathbb{R}^{2 \times 2}.$$

Iterace (10.11) 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. \blacklozenge

Příklad 10.8. 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. \blacklozenge

Příklad 10.9. Pro funkci $g(x) = x^3 - 2x + 2$ zvolme $x_0 = 0$. Další iterace bude $x_1 = 1$ a další $x_2 = 0$. Metoda bude oscilovat mezi hodnotami 0 a 1, tedy bude divergovat. \blacklozenge

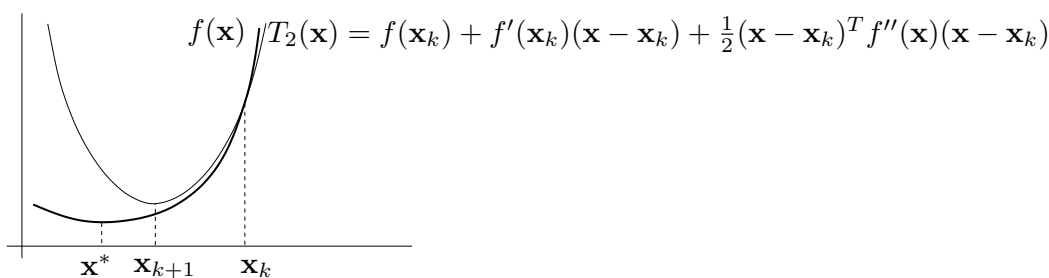
10.4.2 Použití na minimalizaci funkce

Newtonovu metodu lze použít pro hledání lokálního extrému dvakrát diferencovatelné funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tak, že v metodě (10.11) položíme $\mathbf{g}(\mathbf{x}) = f'(\mathbf{x})^T$. Tím dostaneme iteraci

$$\mathbf{x}_{k+1} = \mathbf{x}_k - f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T, \quad (10.12)$$

kde $f''(\mathbf{x}_k)$ je Hessova matice funkce f v bodě \mathbf{x}_k . Neplet' se si tato dvě různá použití Newtonovy metody (tj. na hledání kořenů a na hledání lokálních extrémů)!

Iterace (10.11) se odvodila tak, že se zobrazení \mathbf{g} v okolí bodu \mathbf{x}_k aproximovalo afinním zobrazením $\mathbf{T}_{\mathbf{x}_k}^1$ a pak se našel nulový bod tohoto zobrazení. Lze ukázat (viz Cvičení 10.11), že iterace (10.12) lze odvodit také tak, že se funkce f aproximuje Taylorovým polynomem druhého stupně $T_{\mathbf{x}_k}^2$ (tedy kvadratickou funkcí) a pak se najde minimum této kvadratické funkce.



Iteraci (10.12) lze napsat v obecnějším tvaru (10.1), kde

$$\mathbf{v}_k = -f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T. \quad (10.13)$$

Výhodou tohoto zobecnění je možnost zvolit optimální (ne nutně jednotkovou) délku kroku α_k pomocí jednorozměrné minimalizace (10.2). Metodě (10.12) s jednotkovou délkou kroku (tj. $\alpha_k = 1$ pro každé k) se pak říká **čistá** Newtonova metoda.

Vektoru (10.13) říkáme **Newtonův směr**. Vidíme, že se od gradientního směru (10.5) 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 6.19).

V porovnání s gradientní metodou má Newtonova metoda (použitá na minimalizaci funkce) nevýhodu v tom, že musíme počítat Hessián $f''(\mathbf{x}_k)$ a řešit soustavu $f''(\mathbf{x}_k) \mathbf{v}_k = -f'(\mathbf{x}_k)^T$, což pro velký počet proměnných je pomalé či nemožné. Všimněte si ale, že na rozdíl od §10.4.1 je zde matice $\mathbf{g}'(\mathbf{x}_k) = f''(\mathbf{x}_k)$ symetrická, což může řešení soustavy ulehčit.

10.5 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 přeürčenou, jestliže nemá žádné řešení. Chceme takovou přeürčenou soustavu řešit přibližně ve smyslu nejmenších čtverců. Tedy chceme minimalizovat funkci

$$f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|^2 = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x}) = \sum_{i=1}^m g_i(\mathbf{x})^2, \quad (10.14)$$

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

Zatímco v §10.3 a §10.4.2 bylo cílem minimalizovat *obecnou* funkci, zde chceme minimalizovat funkci ve speciálním tvaru (10.14). Nyní máme dvě možnosti. Bud' můžeme nasadit na funkci (10.14) jednu z metod pro minimalizaci obecné funkce, k čemuž se vrátíme v §10.5.2. Nebo můžeme být chytřejší a využít speciálního tvaru funkce (10.14), což popíšeme v §10.5.1.

10.5.1 Gauss-Newtonova metoda

Aproximujme opět zobrazení \mathbf{g} v okolí bodu \mathbf{x}_k afinním zobrazením $\mathbf{T}_{\mathbf{x}_k}^1$ dle (10.10). Úloha (10.14) pak vyžaduje minimalizovat $\|\mathbf{T}_{\mathbf{x}_k}^1(\mathbf{x})\|^2$. To je úloha lineárních nejmenších čtverců, kterou již známe z §5.1. Normální rovnice (5.3) má tvar

$$\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) = -\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k). \quad (10.15)$$

Její řešení napíšeme pomocí pseudoinverze:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k). \quad (10.16)$$

Metoda (10.16) je známa jako (čistá) **Gauss-Newtonova metoda**. Můžeme ji opět napsat obecněji ve tvaru (10.1) se směrem hledání (Gauss-Newtonův směr)

$$\mathbf{v}_k = -\mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k). \quad (10.17)$$

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 (10.16) se redukuje na Newtonovu metodu (10.11).

Pokud má Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ lineárně nezávislé sloupce (viz §5.1), výraz (10.17) můžeme napsat jako

$$\mathbf{v}_k = -(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k) = -\frac{1}{2}(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)^T, \quad (10.18)$$

kde ve výrazu na pravé straně jsme dosadili derivaci

$$f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x}) \quad (10.19)$$

účelové funkce (10.14) (odvod'te dle §8.5.2!). Vidíme, že Gauss-Newtonův směr (10.18) se liší od gradientního směru (10.5) 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 6.19). 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 6.20), což ovšem již předpokládáme kvůli existenci inverze. Tedy vidíme, že za přirozených podmínek je Gauss-Newtonův směr vždy sestupný.

Čistá Gauss-Newtonova metoda (tj. s jednotkovou délkou kroku) může divergovat, a to i když je počáteční odhad \mathbf{x}_0 libovolně blízko lokálnímu minimu funkce (10.14). Protože ale Gauss-Newtonův směr je vždy sestupný, vhodnou volbou délky kroku α_k lze vždy zajistit konvergenci.

Příklad 10.10. 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 10.7. Ani jeden z těchto průsečíků neleží na třetí křivce (i když je jí blízko), tedy soustava je přeurčená. Nezbyvá nám tedy, než ji řešit přibližně. Hledáme bod $(x, y) \in \mathbb{R}^2$, který minimalizuje číslo

$$f(x, y) = \mathbf{g}(x, y)^T \mathbf{g}(x, y) = ((x-1)^2 + y^2 - 1)^2 + (x^4 + y^4 - 1)^2 + (x^2 + (y-1)^2 - 1/2)^2$$

kde

$$\mathbf{g}(x, y) = \begin{bmatrix} (x-1)^2 + y^2 - 1 \\ x^4 + y^4 - 1 \\ x^2 + (y-1)^2 - 1/2 \end{bmatrix} \in \mathbb{R}^3, \quad \mathbf{g}'(x, y) = \begin{bmatrix} 2(x-1) & 2y \\ 4x^3 & 4y^3 \\ 2x & 2(y-1) \end{bmatrix} \in \mathbb{R}^{3 \times 2}.$$

Rozumný počáteční odhad je $(x_0, y_0) = (1, 1)$. První Gauss-Newtonova iterace (10.16) 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. ♦

Příklad 10.11. 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 přeürčenou nelineární soustavu ve smyslu nejmenších čtverců, tedy minimalizujme funkci

$$f(\mathbf{x}) = \sum_{i=1}^m (\|\mathbf{x} - \mathbf{a}_i\| - y_i)^2.$$

Máme³ 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 (pomůže vám §8.5.2 ale udělejte sami!) $g'_i(\mathbf{x}) = (\mathbf{x} - \mathbf{a}_i)^T / \|\mathbf{x} - \mathbf{a}_i\|$. Tedy

$$\mathbf{g}'(\mathbf{x}) = \begin{bmatrix} (\mathbf{x} - \mathbf{a}_1)^T / \|\mathbf{x} - \mathbf{a}_1\| \\ \vdots \\ (\mathbf{x} - \mathbf{a}_m)^T / \|\mathbf{x} - \mathbf{a}_m\| \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Pak dosadíme do vzorečku (10.16). ♦

10.5.2 Rozdíl oproti Newtonově metodě

Předpokládejme, že bychom optimalizovali naši účelovou funkci (10.14) přímo Newtonovou metodou z §10.4.2. Spočítejme (viz Cvičení 8.12.d) Hessián funkce (10.14):

$$f''(\mathbf{x}) = 2\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x}) + 2 \sum_{i=1}^m g_i(\mathbf{x}) g''_i(\mathbf{x}). \quad (10.20)$$

Hessián je součtem členu obsahujícího derivace prvního řádu a členu obsahujícího derivace druhého řádu. Vidíme, že Gauss-Newtonův směr (10.18) se liší od Newtonova směru (10.13) zanedbáním členu druhého řádu v Hessiánu (10.20). Jinými slovy, Gauss-Newtonovu metodu je možno vnímat jako aproximaci Newtonovy metody na minimalizaci funkce (10.14) spočívající v tom, že zanedbáme členy druhého řádu, tedy skutečný Hessián (10.20) 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 (10.14). Ovšem vyhnuli jsme se počítání druhých derivací funkce \mathbf{g} , což je hlavní výhoda Gauss-Newtonovy metody.

³Zde ignorujeme, že funkce f není všude diferencovatelná. Přesně, není diferencovatelná v bodech $\mathbf{a}_1, \dots, \mathbf{a}_m$ (promyslete!), což budeme ignorovat.

10.5.3 Levenberg-Marquardtova metoda

Levenberg-Marquardtova metoda je široce používané vylepšení Gauss-Newtonovy metody, které její iteraci

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k) \quad (10.21)$$

nahrazuje iterací

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k) + \mu_k \mathbf{I})^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k) \quad (10.22)$$

kde $\mu_k > 0$. Přidání členu $\mu_k \mathbf{I}$ je vlastně (Tichonovova) regularizace (viz (5.18)). Potom:

- Pro malé μ_k se iterace (10.22) blíží Gauss-Newtonově iteraci.
- Pro velké μ_k je $(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k) + \mu_k \mathbf{I})^{-1} \approx \mu_k^{-1} \mathbf{I}$, tedy (10.22) se (po dosazení (10.19)) blíží iteraci $\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{2} \mu_k^{-1} f'(\mathbf{x}_k)^T$ 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 se zvětšujícím se k pomocí jednoduché heuristiky. Začneme s nějakým velkým μ_0 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, třeba 2 nebo 10. Všimněte si, toto nahrazuje optimalizaci délky kroku α_k (*line search*).

Motivaci pro přidání regularizace do Gauss-Newtonovy iterace (10.21) lze vidět i jinak. Matice $\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)$ může být singulární (to když sloupce $\mathbf{g}'(\mathbf{x}_k)$ budou lineárně závislé) nebo blízká singulární. Pak její inverze neexistuje nebo je velmi citlivá na malé změny matice, což může neblaze ovlivnit konvergenci metody. Matice (10.22) je ale vždy pozitivně definitní (viz Cvičení 6.17), a tedy regulární.

10.6 Cvičení

10.1. 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

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

10.2. 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.

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

- 10.3. Najděte lokální extrémy funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ dané vzorcem $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} - \sum_{i=1}^n x_i \log x_i$, kde \mathbf{a} je známý vektor.
- 10.4. Vyšetřete extrémy funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ definované vzorcem $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + 1/(\mathbf{b}^T \mathbf{x})$, kde \mathbf{a} a $\mathbf{b} \neq \mathbf{0}$ jsou známé vektory. Tj. zjistěte, jaké podmínky musí splňovat vektory \mathbf{a} a \mathbf{b} aby funkce měla aspoň jeden extrém a za tohoto předpokladu najděte všechny (lokální i globální) extrémy funkce f .
- 10.5. Najděte všechna řešení rovnice $\sin x = \frac{1}{2}x$ (sinus je v radiánech) na kalkulačce s největší přesností, jakou dokážete.
- 10.6. 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.
- 10.7. 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.
- 10.8. Soustavu dvou rovnic o jedné neznámé

$$x^2 + x = 1$$

$$x^2 - x = 1$$

chceme řešit přibližně ve smyslu nejmenších čtverců. Napište iteraci (a) čisté Gauss-Newtonovy metody, (b) čisté Newtonovy metody. Výsledné vzorce zjednodušte.

- 10.9. Máme soustavu rovnic

$$x + y - 2xy = 1$$

$$-x + y + xy = -3$$

$$x - y + xy = 1$$

Je soustava lineární? Kolik má řešení a proč? Chceme soustavu řešit přibližně ve smyslu nejmenších čtverců, tj. minimalizovat funkci $f(x, y)$ ve tvaru (10.14). Napište iteraci (a) gradientní, (b) Newtonovy, (c) Gauss-Newtonovy, (d) Levenberg-Marquardtovy metody.

- 10.10. Chceme najít vzdálenost množiny $\{(x, y) \in \mathbb{R}^2 \mid x^2 = y\}$ od kružnice s poloměrem 1 a středem v bodě $(2, 0)$. Tvrdíme, že tuto úlohu lze řešit tak, že vyřešíme přeурčenou soustavu $\{x^2 = y, (x - 2)^2 + y^2 = 1\}$ přibližně ve smyslu nejmenších čtverců, tedy minimalizujeme funkci $f(x, y) = (x^2 - y)^2 + ((x - 2)^2 + y^2 - 1)^2$. Je to pravda, bude minimální hodnota této funkce rovna (čtverci) vzdálenosti mezi množinami? Pokud ne, jak bychom tuto vzdálenost spočítali?
- 10.11. Čistá Newtonova metoda (10.12) na minimalizaci funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je Newtonova metoda (10.11) na řešení soustavy $f'(\mathbf{x})^T = \mathbf{0}$. Takto jsme ji odvodili. Ukažte, že iteraci (10.12) 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.
- 10.12. (★) V §10.3.1 jsme ukázali, že iterace gradientní metody není invariantní vůči lineární transformaci souřadnic $\mathbf{y} = \mathbf{A}\mathbf{x}$ (pro regulární \mathbf{A}). Ukažte, že iterace Newtonovy metody (10.12) je invariantní vůči této transformaci.
- 10.13. *Pevný bod* Newtonovy metody (10.11) je takový bod \mathbf{x}_k , který všechny další iterace již nemění, tj. $\mathbf{x}_{k+1} = \mathbf{x}_k$ (z čehož plyne, že $\mathbf{x}_l = \mathbf{x}_k$ pro všechna $l \geq k$). Může mít Newtonova

metoda pevný bod, když soustava $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ nemá řešení (máme na mysli reálné řešení, komplexní mít může)?

- 10.14. (★) Gauss-Newtonův směr (10.17) se získá řešením normální rovnice (10.15). 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 (10.15) má nekonečně mnoho řešení, tj. směrů hledání je nekonečně mnoho. Dokažte, že každý takový směr je sestupný.

Nápověda a řešení

- 10.1.a) funkce nemá v tomto bodě lokální extrém
 10.1.b) nemůžeme rozhodnout, zda má funkce má v tomto bodě lokální extrém
 10.1.c) funkce má v tomto bodě lokální minimum
 10.2.d) Stacionární body jsou 4.
 10.2.e) Stacionární body jsou 3.
 10.2.f) Stacionárních bodů je 5.
 10.2.g) Stacionární body jsou $(0, 0, 0)$ a $(3/2, 3/2, -9/4)$.
- 10.3. Funkce je součtem funkcí jedné proměnné, $f(\mathbf{x}) = \sum_i g_i(x_i)$ kde $g_i(x) = a_i x - x \log x$. Tedy hledání extrémů funkce f se dá převést na nezávislé hledání extrémů funkcí g_i , viz Cvičení 9.11. Přesněji, f bude mít lokální/globální maximum/minimum v bodě $\mathbf{x} = (x_1, \dots, x_n)$, právě když každá g_i bude mít lokální/globální maximum/minimum v bodě x_i . Je $g'_i(x) = a_i - \log x - 1 = 0$, tedy $x = e^{a_i-1}$. Není těžké ověřit, že jde o globální maximum. Shrnutí: f má jediný lokální a zároveň globální extrém, a to maximum v bodě $\mathbf{x} = (e^{a_1-1}, \dots, e^{a_n-1})$.
- 10.4. Stacionární podmínka je $f'(\mathbf{x})^T = \mathbf{a} - \mathbf{b}/(\mathbf{b}^T \mathbf{x})^2 = \mathbf{0}$. Tuto rovnici musíme vyřešit, tj. určit, pro jaká \mathbf{a}, \mathbf{b} má řešení a jaká je v tom případě její množina řešení. Nemůže být $\mathbf{a} = \mathbf{b} = \mathbf{0}$, protože $\mathbf{b} = \mathbf{0}$ je zakázáno v zadání (jinak by f nebyla definována). Jistě musí být $\mathbf{b}^T \mathbf{x} \neq \mathbf{0}$, protože jinak by také $f(\mathbf{x})$ nebylo definováno. Pro každé $\alpha \geq 0$ existuje \mathbf{x} takové, že $(\mathbf{b}^T \mathbf{x})^2 = \alpha$ (proč?). Proto má rovnice řešení právě tehdy, když $\mathbf{b} \neq \mathbf{0}$ a existuje skalár $\alpha > 0$ takový, že $\mathbf{a} = \mathbf{b}/\alpha$. Neboli vektory \mathbf{a}, \mathbf{b} jsou nenulové, rovnoběžné a mají stejný směr. (Tento výsledek je intuitivně přijatelný, představíme-li $n = 2$. Pro $\mathbf{a} = (1, 0)$ a $\mathbf{b} = (0, 1)$ (tj. nejsou rovnoběžné) dostaneme funkci $f(x, y) = x + 1/y$, která je očividně neomezená. Pro $\mathbf{a} = (1, 0)$ a $\mathbf{b} = -\mathbf{a}$ (jsou rovnoběžné ale mají opačný směr) dostaneme $f(x, y) = x - 1/x$, a ta také neomezená (načrtněte si graf).)
- Za této podmínky můžeme naši funkci napsat jako $f(\mathbf{x}) = \mathbf{b}^T \mathbf{x}/\alpha + 1/(\mathbf{b}^T \mathbf{x})$, kde $\mathbf{b} \in \mathbb{R}^n$ a $\alpha > 0$ jsou známé. Tato funkce závisí jen na součinu $\mathbf{b}^T \mathbf{x}$ (můžeme se totiž pohybovat jen po přímce dané společným směrem vektorů \mathbf{a}, \mathbf{b}). Označíme-li $\mathbf{b}^T \mathbf{x} = y$, je $f(\mathbf{x}) = g(\mathbf{b}^T \mathbf{x})$ kde $g(y) = y/\alpha + 1/y$. Nyní stačí najít extrémy funkce g . Stacionární podmínka je $g'(y) = 1/\alpha - 1/y^2 = 0$. Tato rovnice má dvě řešení $y = \pm\alpha$. Pomocí jednoduchých úvah (načrtnutí grafu, druhá derivace) zjistíme, že kladný kořen je globální minimum a záporný je globální maximum.
- Odpověď na otázku v zadání: funkce f má aspoň jeden lokální extrém právě tehdy, když $\mathbf{b} \neq \mathbf{0}$ a existuje skalár $\alpha > 0$ takový, že $\mathbf{a} = \mathbf{b}/\alpha$. Za této podmínky má funkce f lokální a zároveň globální minimum v bodech \mathbf{x} splňujících $\mathbf{b}^T \mathbf{x} = \alpha$ a lokální a zároveň globální maximum v bodech \mathbf{x} splňujících $\mathbf{b}^T \mathbf{x} = -\alpha$.
- 10.5. Jeden kořen je $x = 0$ a pak dva další lišící se znaménkem. Jeden z nich získáme Newtonovou metodou: $x_{k+1} = x_k - (2 \sin x_k - x_k)/(2 \cos x_k - 1)$. Načrtneme si grafy funkcí $\sin x$ a $1/\overline{2x}$ a z toho zvolíme počáteční odhad, např. $x_0 = 2$. Po několika iteracích máme $x_k = 1.895494267033981$.

- 10.8. (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
- $$x \leftarrow x - (\mathbf{g}'(x)^T \mathbf{g}'(x))^{-1} \mathbf{g}'(x)^T \mathbf{g}(x) = x - (8x^2 + 2)^{-1} \begin{bmatrix} 2x + 1 & 2x - 1 \end{bmatrix} \begin{bmatrix} x^2 + x - 1 \\ x^2 - x - 1 \end{bmatrix} = \frac{2x(x^2 + 1)}{4x^2 + 1}.$$
- (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.
- 10.9. 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.
- 10.10. Nebude, vzdálenost by se musela počítat jako minimalizace $(x-u)^2 + (y-v)^2$ za podmínek $x^2 = y$ a $(u-2)^2 + v^2 = 1$. Protože se křivky neprotínají, tato formulace jde zjednodušit: vzdálenost bodu od kružnice je rovna vzdálenost bodu od středu kružnice, tedy stačí minimalizovat $(x-2)^2 + y^2$ za podmínky $x^2 = y$, což se zjednoduší na minimalizaci funkce jedné proměnné $(x-2)^2 + x^4$.
- 10.11. Po přejmenování proměnných $\mathbf{x}_k, \mathbf{x}_{k+1}$ na \mathbf{x}, \mathbf{y} máme ukázat, že stacionární bod \mathbf{y} Taylorova polynomu (8.23c) 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 polynomu (s užitím Cvičení 8.12).
- 10.13. Aby $\mathbf{x}_{k+1} = \mathbf{x}_k$, musí být $\mathbf{g}'(\mathbf{x}_k)^{-1} \mathbf{g}(\mathbf{x}_k) = \mathbf{0}$. Matice $\mathbf{g}'(\mathbf{x}_k)$ je regulární (protože předpokládáme, že má inverzi) a tedy i matice $\mathbf{g}'(\mathbf{x}_k)^{-1}$ je regulární. Ale nemůže být $\mathbf{g}(\mathbf{x}_k) = \mathbf{0}$, protože soustava nemá řešení. Tedy pevný bod neexistuje.
- 10.14. Pišme normální rovnici (10.15) 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 6.26), 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 11

Lokální extrémý vázané rovnostmi

Hledejme lokální extrémý funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \}, \quad (11.1)$$

kde $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$ je zobrazení se složkami g_1, \dots, g_m . To odpovídá úloze (1.10) (nebo její maximalizační verzi) s omezeními typu rovnosti:

$$\begin{aligned} & \min/\max \quad f(x_1, \dots, x_n) \\ & \text{za podmíněk} \quad g_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, m. \end{aligned} \quad (11.2)$$

Mluvíme o extréměch 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 podmínky na volné lokální extrémý ze §10.1 (viz Tvzení 9.3). Někdy ovšem lze vyjádřit všechna řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ parametricky (neboli parametrizovat množinu X) a úlohu tak převést na úlohu bez omezení. To jsme použili v Příkladu 1.2, zde je další příklad:

Příklad 11.1. Řešme úlohu

$$\begin{aligned} & \max \quad x_1 + x_2 \\ & \text{za podmínky} \quad x_1^2 + x_2^2 = 1 \end{aligned} \quad (11.3)$$

tedy máme $m = 1$ a $n = 2$ a maximalizujeme účelovou funkci $f(x_1, x_2) = x_1 + x_2$ na kružnici $X = \{ (x_1, x_2) \in \mathbb{R}^2 \mid g(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0 \}$. Množinu X lze parametrizovat jako

$$X = \{ (\cos \alpha, \sin \alpha) \mid \alpha \in \mathbb{R} \}, \quad (11.4)$$

což převede úlohu na maximalizaci funkce $\varphi(\alpha) = f(\cos \alpha, \sin \alpha) = \cos \alpha + \sin \alpha$ na množině \mathbb{R} . Je-li α lokální extrém funkce φ , pak dle Věty 10.3 je $\varphi'(\alpha) = -\sin \alpha + \cos \alpha = 0$, což má řešení $\alpha = \frac{\pi}{4} + k\pi$. To odpovídá bodům $(x_1, x_2) = \pm \frac{\sqrt{2}}{2}(1, 1)$. To jsou tedy body podezřelé z lokálního extrému. \blacklozen

Někdy ovšem množinu (11.1) parametrizovat nelze, nebo je to složité nebo nevýhodné.

Příklad 11.2. Zobecněme úlohu (11.3): pro dané $\mathbf{a} \in \mathbb{R}^n$ řešme

$$\begin{aligned} & \max \quad \mathbf{a}^T \mathbf{x} \\ & \text{za podmínky} \quad \mathbf{x}^T \mathbf{x} = 1 \end{aligned} \quad (11.5)$$

Maximalizujeme lineární funkci $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x}$ na množině $X = \{ \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = \mathbf{x}^T \mathbf{x} - 1 = 0 \}$, což je n -rozměrná sféra. Zde už není vůbec jasné, jak bychom mohli parametrizovat množinu X podobně jako v (11.4). Přitom řešení úlohy (11.5) po krátkém zamyšlení uhodnete (nakreslete si obrázky pro $n = 1, 2, 3$): maximum se nabývá pro $\mathbf{x} = \mathbf{a}/\|\mathbf{a}\|$.

Asi vás napadne vyjádřit jednu proměnnou (např. x_n) z podmínky $\mathbf{x}^T \mathbf{x} = x_1^2 + \dots + x_n^2 = 1$ a dosadit ji do účelové funkce $\mathbf{a}^T \mathbf{x} = a_1 x_1 + \dots + a_n x_n$, čímž dostaneme úlohu bez omezení. Musíme ale uvažovat zvlášť dva případy $x_n = \pm \sqrt{1 - x_1^2 - \dots - x_{n-1}^2}$. Tento způsob řešení je složitý a nehezky, už proto že nás donutil pracovat se složkami vektorů \mathbf{a}, \mathbf{x} (téměř vždy je lépe chovat se k vektorům a maticím jako k nedělitelným objektům). ♦

Dále uvedeme jiné podmínky na lokální extrémů vázané rovnostmi, vyjádřené s pomocí jistých pomocných proměnných, tzv. *Lagrangeových multiplikátorů*.

11.1 Lineární omezení

Uvažujme nejprve důležitý speciální případ, kdy zobrazení \mathbf{g} je afinní, tj. $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ kde $\mathbf{A} \in \mathbb{R}^{m \times n}$. Dle Věty 3.18 je tedy množina X afinní podprostor \mathbb{R}^n . Úlohu můžeme psát jako

$$\min/\max \{ f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b} \}, \quad (11.6)$$

tj. minimalizujeme nebo maximalizujeme funkci f za podmínek lineárních rovností. Větu 10.2 snadno zobecníme pro tento případ:

Věta 11.1. *Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \text{null } \mathbf{A}$.*

- *Jestliže \mathbf{x} je lokální minimum funkce f vázané podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $f_{\mathbf{v}}(\mathbf{x}) \geq 0$.*
- *Jestliže \mathbf{x} je lokální maximum funkce f vázané podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $f_{\mathbf{v}}(\mathbf{x}) \leq 0$.*

Důkaz. Necht' $\mathbf{A}\mathbf{x} = \mathbf{b}$ a $\mathbf{v} \in \text{null } \mathbf{A}$ (tj. $\mathbf{A}\mathbf{v} = \mathbf{0}$). Pak pro každé $\alpha \in \mathbb{R}$ je $\mathbf{A}(\mathbf{x} + \alpha\mathbf{v}) = \mathbf{b}$, tedy $\mathbf{x} + \alpha\mathbf{v} \in X$. Jestliže $f_{\mathbf{v}}(\mathbf{x}) < 0$, pak tedy dle Tvzení 10.1 pro každé $\varepsilon > 0$ existuje $\alpha > 0$ tak, že $\mathbf{x} + \alpha\mathbf{v} \in X \cap B_{\varepsilon}(\mathbf{x})$ a $f(\mathbf{x} + \alpha\mathbf{v}) < f(\mathbf{x})$. Tedy \mathbf{x} není lokální minimum funkce f na množině X (viz definice lokálního minima v §9.3). Obdobně pro maximum. ■

Důkaz má jasný geometrický význam: jestliže v nějakém bodě je směr $\mathbf{v} \in \text{null } \mathbf{A}$ sestupný [vzestupný] pro f , pak hodnotu f můžeme zmenšit [zvětšit] malým posunutím bodu ve směru \mathbf{v} , čímž neopustíme afinní podprostor X . Proto bod nemůže být lokální minimum [maximum].

Pro (totálně) diferencovatelnou funkci f zobecníme Větu 10.3:

Věta 11.2. *Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je (totálně) diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$.*

Jestliže \mathbf{x} je lokální extrém funkce f vázaný podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $\nabla f(\mathbf{x}) \perp \text{null } \mathbf{A}$.

Důkaz. Dle Věty 8.5 je $f_{\mathbf{v}}(\mathbf{x}) = f'(\mathbf{x})\mathbf{v}$. Protože $\text{null } \mathbf{A}$ je podprostor, $\mathbf{v} \in \text{null } \mathbf{A}$ platí právě když $-\mathbf{v} \in \text{null } \mathbf{A}$. Jestliže \mathbf{x} je lokální minimum f vázané podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak dle Věty 11.1 pro každé $\mathbf{v} \in \text{null } \mathbf{A}$ platí $f_{\mathbf{v}}(\mathbf{x}) = f'(\mathbf{x})\mathbf{v} \geq 0$ a $f_{-\mathbf{v}}(\mathbf{x}) = -f'(\mathbf{x})\mathbf{v} \geq 0$, tedy $f'(\mathbf{x})\mathbf{v} = 0$. To znamená, že vektor $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ je kolmý na podprostor $\text{null } \mathbf{A}$. ■

Uvedeme ještě jiný důkaz Věty 11.2 (který nám umožní dokázat podmínky druhého řádu):

Důkaz. Množinu řešení soustavy $\mathbf{Ax} = \mathbf{b}$ parametrizujeme jako $\mathbf{x} = \mathbf{By} + \mathbf{x}_0$, kde \mathbf{x}_0 a \mathbf{B} jsou libovolné splňující $\mathbf{Ax}_0 = \mathbf{b}$ a $\text{rng } \mathbf{B} = \text{null } \mathbf{A}$. Naše úloha je tedy ekvivalentní hledání lokálních extrémů funkce

$$\varphi(\mathbf{y}) = f(\mathbf{By} + \mathbf{x}_0) \quad (11.7)$$

bez omezení. Dle Věty 10.3 lokální extrémy \mathbf{y} splňují

$$\varphi'(\mathbf{y}) = f'(\mathbf{By} + \mathbf{x}_0)\mathbf{B} = f'(\mathbf{x})\mathbf{B} = \mathbf{0}, \quad (11.8)$$

kde jsme užili řetízkové pravidlo (jako v Příkladě 8.13). To říká, že gradient $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ je kolmý na podprostor $\text{rng } \mathbf{B} = \text{null } \mathbf{A}$. ■

Podmínka $\nabla f(\mathbf{x}) \perp \text{null } \mathbf{A}$ ve Větě 11.2 znamená, že gradient $\nabla f(\mathbf{x})$ je kolmý na afinní podprostor X . Lze ji podrobněji napsat jako (viz Věta 4.8)

$$\nabla f(\mathbf{x}) \in (\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T). \quad (11.9)$$

To říká, že vektor $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ je lineární kombinací řádků matice \mathbf{A} , neboli $f'(\mathbf{x}) = \boldsymbol{\lambda}^T \mathbf{A}$ pro nějaké $\boldsymbol{\lambda} \in \mathbb{R}^m$. Tedy pro lokální extrém platí

$$f'(\mathbf{x}) = \boldsymbol{\lambda}^T \mathbf{A}, \quad (11.10a)$$

$$\mathbf{Ax} = \mathbf{b}. \quad (11.10b)$$

Tato soustava má $m + n$ rovnic a $m + n$ neznámých. Později řekneme, že prvky vektoru $\boldsymbol{\lambda}$ se nazývají *Lagrangeovy multiplikátory*.

Příklad 11.3. Vraťme se k úloze (5.19), tedy k hledání řešení nehomogenní lineární soustavy s nejmenší normou. Místo funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$ budeme minimalizovat funkci $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{x}$, což úlohu nezmění a stacionární podmínky nám vyjdou hezčí. Je $f'(\mathbf{x}) = \mathbf{x}^T$, tedy rovnost (11.10a) je $\mathbf{x} = \mathbf{A}^T \boldsymbol{\lambda}$. Soustava (11.10) je tedy soustava (5.21), kterou jsme v §5.2 odvodili úvahou. ◆

Příklad 11.4. Obecněji, řešme úlohu

$$\begin{aligned} \min \quad & \|\mathbf{Ax} - \mathbf{b}\|^2 \\ \text{za podmínek} \quad & \mathbf{Cx} = \mathbf{d} \end{aligned} \quad (11.11)$$

Tedy řešíme úlohu nejmenších čtverců (5.2) s lineárními omezeními (angl. *linearly constrained least squares*). Tato úloha má hodně aplikací.

Máme (opět po přidání $\frac{1}{2}$ pro pohodlí)

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|^2 = \frac{1}{2} (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}), \\ f'(\mathbf{x}) &= \mathbf{x}^T \mathbf{A}^T \mathbf{A} - \mathbf{b}^T \mathbf{A} \end{aligned}$$

(derivaci spočtete sami!). Tedy podmínka stacionarity (11.10) je

$$\mathbf{A}^T \mathbf{Ax} + \mathbf{C}^T \boldsymbol{\lambda} = \mathbf{A}^T \mathbf{b}, \quad (11.12a)$$

$$\mathbf{Cx} = \mathbf{d}, \quad (11.12b)$$

neboli

$$\begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T \mathbf{b} \\ \mathbf{d} \end{bmatrix}.$$

Tuto lineární soustavu vyřešíme jednou ze známých metod (k tomu viz Cvičení 11.24). ◆

Věta 11.2 udává nutnou podmínku prvního řádu na lokální extrém vázaný lineárními rovnostmi. Uvedeme nyní podmínky druhého řádu. K tomu potřebujeme nový pojem: matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ je *pozitivně semidefinitní na podprostoru* $Y \subseteq \mathbb{R}^n$, jestliže $\mathbf{x}^T \mathbf{C} \mathbf{x} \geq 0$ pro každé $\mathbf{x} \in Y$. Jak tuto podmínku ověříme? Najdeme matici \mathbf{B} tak, že $Y = \text{rng } \mathbf{B}$ (např. sloupce \mathbf{B} tvoří bázi podprostoru Y). Pak $\mathbf{x} \in Y$ právě když $\mathbf{x} = \mathbf{B}\mathbf{y}$ pro nějaké \mathbf{y} . Protože $\mathbf{x}^T \mathbf{C} \mathbf{x} = \mathbf{y}^T \mathbf{B}^T \mathbf{C} \mathbf{B} \mathbf{y}$, převedli jsme problém na ověřování pozitivní semidefinitnosti matice $\mathbf{B}^T \mathbf{C} \mathbf{B}$. Podobně definujeme pozitivní a negativní (semi)definitnost a indefinitnost matice na podprostoru.

Věta 11.3. *Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je dvakrát diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$.*

- *Jestliže \mathbf{x} je lokální minimum [maximum] funkce f vázané podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak existuje λ tak že (\mathbf{x}, λ) splňují soustavu (11.10) a Hessova matice $f''(\mathbf{x})$ je pozitivně [negativně] semidefinitní na podprostoru $\text{null } \mathbf{A}$.*
- *Jestliže existuje λ tak že (\mathbf{x}, λ) splňují soustavu (11.10) a Hessova matice $f''(\mathbf{x})$ je pozitivně [negativně] definitní na podprostoru $\text{null } \mathbf{A}$, pak \mathbf{x} je ostré lokální minimum [maximum] funkce f vázané podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$.*

Důkaz. V druhém důkazu Věty 11.2 jsme ukázali, že $\mathbf{x} = \mathbf{B}\mathbf{y} + \mathbf{x}_0$ je lokální extrém funkce f vázaný podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, právě když \mathbf{y} je lokální extrém funkce φ bez omezení. Je (viz §8.8)

$$\varphi''(\mathbf{y}) = \mathbf{B}^T f''(\mathbf{B}\mathbf{y} + \mathbf{x}_0) \mathbf{B} = \mathbf{B}^T f''(\mathbf{x}) \mathbf{B}.$$

Zbytek plyne z podmínek druhého řádu ve Větě 10.4, použité na funkci φ . ■

V Příkladech 11.3 a 11.4 byl Hessián $f''(\mathbf{x})$ pozitivně semidefinitní, tedy byl pozitivně semidefinitní i na každém podprostoru. Zde je příklad, kdy tomu tak není:

Příklad 11.5. Řešme úlohu

$$\begin{aligned} \max \quad & x_1 x_2 + x_2 x_3 + x_1 x_3 \\ \text{za podmínky} \quad & x_1 + x_2 + x_3 = 3. \end{aligned} \tag{11.13}$$

Řešení podmínek prvního řádu (11.10) je $x_1 = x_2 = x_3 = 1$, $\lambda = 2$. Hessián

$$f''(x_1, x_2, x_3) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

je indefinitní (na \mathbb{R}^3). Zjistíme jeho definitnost na podprostoru $\text{null } \mathbf{A}$. Máme

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{B}^T f''(x_1, x_2, x_3) \mathbf{B} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

kde sloupce matice \mathbf{B} jsou (libovolná) báze podprostoru $\text{null } \mathbf{A}$. Poslední matice je negativně definitní, tedy $f''(x_1, x_2, x_3)$ je negativně definitní na podprostoru $\text{null } \mathbf{A}$, tedy $(x_1, x_2, x_3) = (1, 1, 1)$ je (ostré) lokální maximum úlohy. ◆

11.2 Nelineární omezení

Přejděme nyní k obecnému případu, kdy \mathbf{g} je libovolné (ne nutně afinní) diferencovatelné zobrazení. Podmínka prvního řádu na lokální extrémy pro tato omezení bude podobná jako pro lineární omezení, ale na rozdíl od Věty 11.2 ji nedokážeme, protože důkaz by byl příliš dlouhý. Pouze vysvětlíme geometrický význam této podmínky a uvedeme příklady.

11.2.1 Tečný prostor

Je-li zobrazení \mathbf{g} v nějakém bodě \mathbf{x} diferencovatelné, můžeme ho v okolí bodu \mathbf{x} aproximovat jeho Taylorovým polynomem prvního stupně (8.24)

$$\mathbf{g}(\mathbf{y}) \approx \mathbf{T}_{\mathbf{x}}^1(\mathbf{y}) = \mathbf{g}(\mathbf{x}) + \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (11.14)$$

kde $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ neboť $\mathbf{x} \in X$. Množina X se tím změní na

$$\{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{0}\} = \{\mathbf{x} + \mathbf{y}' \mid \mathbf{y}' \in \mathbb{R}^n, \mathbf{g}'(\mathbf{x})\mathbf{y}' = \mathbf{0}\} = \mathbf{x} + \text{null } \mathbf{g}'(\mathbf{x}), \quad (11.15)$$

kde jsme udělali substituci $\mathbf{y}' = \mathbf{y} - \mathbf{x}$. Množina (11.15) je afinní podprostor, je to (lineární) podprostor $\text{null } \mathbf{g}'(\mathbf{x})$ posunutý¹ do bodu \mathbf{x} (viz §3.3). Protože z definice totální derivace (viz §8.5) se zobrazení \mathbf{g} v okolí bodu \mathbf{x} podobá afinnímu zobrazení (11.14), mohli bychom doufat, že množina X se v okolí bodu \mathbf{x} bude podobat afinnímu podprostoru (11.15). Tak tomu ale překvapivě být nemusí (jak uvidíme na příkladech), je k tomu třeba následující dodatečná podmínka.

Bod $\mathbf{x} \in \mathbb{R}^n$ nazveme **regulární bod** zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, jestliže zobrazení \mathbf{g} je v bodě \mathbf{x} spojitě diferencovatelné a Jacobiho matice $\mathbf{g}'(\mathbf{x})$ má lineárně nezávislé řádky (tj. hodnost m). Připomeňme (viz §8.5), že řádky matice $\mathbf{g}'(\mathbf{x})$ jsou transponované gradienty $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$ složek zobrazení \mathbf{g} v bodě \mathbf{x} .

Pozorování 11.4. Je-li $\mathbf{x} \in X$ regulární bod zobrazení \mathbf{g} , pak množina X je v okolí bodu \mathbf{x} podobná afinnímu podprostoru

$$\mathbf{x} + \text{null } \mathbf{g}'(\mathbf{x}) \quad (11.16)$$

dimenze $n - m$. Tento podprostor je tečný k množině X v bodě \mathbf{x} .

Tento fakt uvádíme jen jako neformální pozorování², pro jeho formalizaci bychom předtím museli definovat pojmy ‘podobná’ a ‘tečný’. Jeho význam intuitivně pochopíte na příkladech. Obvykle se za **tečný prostor** k množině X v bodě \mathbf{x} považuje přímo prostor $\text{null } \mathbf{g}'(\mathbf{x})$ a posunutí \mathbf{x} se neuvádí (protože každému je jasné, že tečný prostor prochází bodem \mathbf{x}).

Příklad 11.6. Necht'

$$g: \mathbb{R}^n \rightarrow \mathbb{R}, \quad g(\mathbf{x}) = \mathbf{x}^T \mathbf{x} - 1. \quad (11.17)$$

Množina X je jednotková n -rozměrná sféra. Pro každé $\mathbf{x} \in X$ je $g'(\mathbf{x})^T = \nabla g(\mathbf{x}) = 2\mathbf{x} \neq \mathbf{0}$, tedy každý bod na sféře je regulární bod³ funkce g . Tečný prostor k množině X v bodě $\mathbf{x} \in X$ je nadrovina $\text{null}(\mathbf{x}^T) = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{y} = 0\}$. Posuneme-li ho do bodu \mathbf{x} , je to nadrovina $\mathbf{x} + \text{null}(\mathbf{x}^T) = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{y} = 1\}$ (odvod'te dle (11.15)!).

Pro $n = 1$ množina $X = \{-1, 1\}$ obsahuje jen dva body a tečný prostor v kterémkoliv z těchto bodů je tento bod sám. Pro $n = 2$ je množina X kružnice v \mathbb{R}^2 a tečný prostor v bodě $\mathbf{x} = (x_1, x_2) \in X$ je tečna k této kružnici. Pro $n = 3$ je množina X obyčejná sféra v \mathbb{R}^3 a tečný prostor v bodě $\mathbf{x} = (x_1, x_2, x_3) \in X$ je tečná rovina ke sféře v tomto bodě. ♦

¹Zde by vám mělo být jasné, že ‘+’ ve výrazu $\mathbf{x} + \text{null } \mathbf{g}'(\mathbf{x})$ označuje operaci definovanou v (3.33).

²To, že množina X je v okolí bodu $\mathbf{x} \in X$ ‘podobná’ afinnímu podprostoru, intuitivně znamená, že malinký mravenec lezoucí po množině X v blízkosti bodu \mathbf{x} by nerozlišil, zda leze po (‘zakřivené’) množině X nebo po ‘plochem’ afinním podprostoru (11.16). Jsou-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 *diferenciální geometrie*.

³Uvědomte si, že lineární nezávislost řádků matice $g'(\mathbf{x}) \in \mathbb{R}^{1 \times m}$ znamená, že její jediný řádek je nenulový.

Příklad 11.7. Necht'

$$\mathbf{g}: \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad \mathbf{g}(x_1, x_2, x_3) = (x_1^2 + x_2^2 + x_3^2 - 1, (x_1 - 1)^2 + x_2^2 + x_3^2 - 1).$$

Množina X je průnik dvou jednotkových sfér v \mathbb{R}^3 se středy $(0, 0, 0)$ a $(1, 0, 0)$. Tento průnik je kružnice v \mathbb{R}^3 . Z geometrie je vidět (algebraicky to nebudeme dokazovat), že pro každý bod $(x_1, x_2, x_3) \in X$ jsou vektory $\nabla g_1(x_1, x_2, x_3) = 2(x_1, x_2, x_3)$ a $\nabla g_2(x_1, x_2, x_3) = 2(x_1 - 1, x_2, x_3)$ lineárně nezávislé, tedy všechny body na X jsou regulární body zobrazení \mathbf{g} . Tečný prostor k množině X v bodě $(x_1, x_2, x_3) \in X$ je přímka

$$\text{null} \begin{bmatrix} x_1 & x_2 & x_3 \\ x_1 - 1 & x_2 & x_3 \end{bmatrix} = \text{span}\{(0, x_3, -x_2)\}$$

tečná ke kružnici. Posunutá do bodu (x_1, x_2, x_3) je tato přímka afinní podprostor

$$(x_1, x_2, x_3) + \text{span}\{(0, x_3, -x_2)\} = \{(x_1, x_2 + \alpha x_3, x_3 - \alpha x_2) \mid \alpha \in \mathbb{R}\}.$$

Tato přímka je průnik tečných rovin k oběma sférám v bodě (x_1, x_2, x_3) . \blacklozenge

Ukažme nyní příklady, kdy bod $\mathbf{x} \in X$ není regulární bod zobrazení \mathbf{g} .

Příklad 11.8. Necht'

$$\mathbf{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \mathbf{g}(x_1, x_2) = ((x_1 + 1)^2 + x_2^2 - 1, (x_1 - 1)^2 + x_2^2 - 1). \quad (11.18)$$

Množina X je průnik dvou jednotkových kružnic se středy $(-1, 0)$ a $(1, 0)$. Ty se protínají v jediném bodě $(x_1, x_2) = (0, 0)$. V tomto bodě jsou vektory $\nabla g_1(x_1, x_2) = 2(x_1 + 1, x_2) = (2, 0)$ a $\nabla g_2(x_1, x_2) = 2(x_1 - 1, x_2) = (-2, 0)$ lineárně závislé, tedy to není regulární bod zobrazení \mathbf{g} . Množina X obsahuje jediný bod, tedy je to afinní podprostor dimenze $n - m = 0$. Přesto podprostor $\text{null } \mathbf{g}'(x_1, x_2) = \text{span}\{(0, 1)\}$ není tečný k množině X (tečna k bodu je bod sám). \blacklozenge

Příklad 11.9. Necht'

$$g: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g(x_1, x_2) = x_1 x_2. \quad (11.19)$$

Množina X je sjednocení svislé a vodorovné osy. Bod $(x_1, x_2) = (0, 0)$ není regulární pro funkci g , neboť $\nabla g(x_1, x_2) = (x_2, x_1) = (0, 0)$. Množina X se v okolí tohoto bodu nepodobá žádnému afinnímu podprostoru. Prostor $\text{null} [0 \ 0] = \mathbb{R}^2$ zjevně není tečný k množině X v bodě $(0, 0)$. \blacklozenge

Může se také stát, že množina X je v okolí nějakého bodu $\mathbf{x} \in X$ podobná afinnímu podprostoru dimenze $n - m$, ale přesto tento bod není regulární a tedy prostor (11.16) v tom bodě není tečný k X .

Příklad 11.10. Necht' $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je funkce

$$g(x, y) = (x_1^2 + x_2^2 - 1)^2. \quad (11.20)$$

Protože pro každé číslo $z \in \mathbb{R}$ platí $z = 0 \Leftrightarrow z^2 = 0$, je

$$X = \{(x_1, x_2) \mid (x_1^2 + x_2^2 - 1)^2 = 0\} = \{(x_1, x_2) \mid x_1^2 + x_2^2 - 1 = 0\}.$$

Tedy množina X je kružnice (jako v Příkladě 11.6 pro $n = 2$). Všimněte si důležité věci: různá zobrazení \mathbf{g} mohou definovat stejnou množinu X (proto jsme regularitu bodu definovali vzhledem k zobrazení \mathbf{g} a ne vzhledem k množině X). Máme $\nabla g(x_1, x_2) = 4(x_1^2 + x_2^2 - 1)(x_1, x_2)$. Pro každý bod $(x, y) \in X$ je $\nabla g(x_1, x_2) = (0, 0)$, tedy bod (x_1, x_2) není regulární pro funkci g . Prostor $\text{null} [0 \ 0] = \mathbb{R}^2$ zjevně nemá tečný ke kružnici. \blacklozenge

Příklad 11.11. Speciálně, pro $m = 1$ je množina X vrstevnice funkce g nulové výšky a $g'(\mathbf{x})^T = \nabla g(\mathbf{x})$. Dostali jsme tedy tvrzení z §8.7, že gradient je kolmý k vrstevnici. \blacklozenge

11.2.2 Podmínka prvního řádu

S intuicí osvojenou v §11.2.1 by vás nyní nemělo překvapit, že Větu 11.2 lze zobecnit na (ne nutně afinní) diferencovatelná zobrazení \mathbf{g} tak, že podprostor $\text{null } \mathbf{A}$ nahradíme tečným podprostorem $\text{null } \mathbf{g}'(\mathbf{x})$ k množině X v bodě lokálního extrému \mathbf{x} . Větu uvádíme bez důkazu:

Věta 11.5. *Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je (totálně) diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$.
Necht' \mathbf{x} je regulární bod zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$.
Jestliže \mathbf{x} je lokální extrém funkce f vázaný podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, pak $\nabla f(\mathbf{x}) \perp \text{null } \mathbf{g}'(\mathbf{x})$.*

Podmínka $\nabla f(\mathbf{x}) \perp \text{null } \mathbf{g}'(\mathbf{x})$ říká, že gradient $\nabla f(\mathbf{x})$ je kolmý k tečnému prostoru k množině X (tedy vlastně k množině X) v bodě \mathbf{x} . Lze ji podrobněji napsat jako (viz Věta 4.8)

$$\nabla f(\mathbf{x}) \in (\text{null } \mathbf{g}'(\mathbf{x}))^\perp = \text{rng}(\mathbf{g}'(\mathbf{x})^T) = \text{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\}, \quad (11.21)$$

tedy gradient $\nabla f(\mathbf{x}) = f'(\mathbf{x})$ je lineární kombinací řádků matice $\mathbf{g}'(\mathbf{x})$. Lokální extrémy f vázaný podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ tedy splňují (za podmínek Věty 11.5)

$$f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x}) = \mathbf{0}, \quad (11.22a)$$

$$\mathbf{g}(\mathbf{x}) = \mathbf{0} \quad (11.22b)$$

pro nějaká čísla $(\lambda_1, \dots, \lambda_m) = \boldsymbol{\lambda} \in \mathbb{R}^m$ (tzv. **Lagrangeovy multiplikátory**). Soustava (11.22) má $m + n$ rovnic a stejný počet neznámých.

Soustava (11.22) se často zapisuje pomocí **Lagrangeovy funkce** $L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ s hodnotami

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \dots + \lambda_m g_m(\mathbf{x}). \quad (11.23)$$

Rovnici (11.22a) pak lze psát jako $L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$ a⁴ rovnici (11.22b) jako $L_{\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{g}(\mathbf{x})^T = \mathbf{0}$. Tedy řešení $(\mathbf{x}, \boldsymbol{\lambda})$ soustavy (11.22) jsou stacionární bod Lagrangeovy funkce, tj. soustavu (11.22) lze psát jako $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$.

Příklad 11.12. Řešme znovu Příklad 11.1. 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

$$L_x(x, y, \lambda) = 1 - 2\lambda x = 0,$$

$$L_y(x, y, \lambda) = 1 - 2\lambda y = 0,$$

$$L_\lambda(x, y, \lambda) = 1 - x^2 - y^2 = 0.$$

Kvůli podmínce $x^2 + y^2 = 1$ nemůžou být x ani y nulové, tedy (z prvních dvou rovnic) ani λ ne. 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, $(x, y) = \pm(1, 1)/\sqrt{2}$. \blacklozenge

⁴Výraz $L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda})$ označuje řádkový vektor parciálních derivací funkce L podle x_1, \dots, x_n . Podobně $L_{\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda})$.

Příklad 11.13. Řešme Příklad 11.1, kde ale omezení změňme na $g(x, y) = (1 - x^2 - y^2)^2 = 0$. Podle Příkladu 11.10 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

$$\begin{aligned} L_x(x, y, \lambda) &= 1 - 4\lambda x(1 - x^2 - y^2) = 0, \\ L_y(x, y, \lambda) &= 1 - 4\lambda y(1 - x^2 - y^2) = 0, \\ L_\lambda(x, y, \lambda) &= (1 - x^2 - y^2)^2 = 0. \end{aligned}$$

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. ♦

Příklad 11.14. Získejme podmínky stacionarity (11.10) pro úlohu (11.6) s lineárními omezeními pomocí formalismu s Lagrangeovou funkcí. Máme

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T(\mathbf{b} - \mathbf{A}\mathbf{x})$$

tedy

$$\begin{aligned} L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) &= f'(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{A} = \mathbf{0}, \\ L_{\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) &= \mathbf{b} - \mathbf{A}\mathbf{x} = \mathbf{0}. \end{aligned}$$

Příklad 11.15. Vraťme se k úloze (11.5). Máme

$$L(\mathbf{x}, \lambda) = \mathbf{a}^T \mathbf{x} + \frac{1}{2}\lambda(1 - \mathbf{x}^T \mathbf{x}),$$

kde jsme pro pohodlí napsali $\frac{1}{2}\lambda$ místo λ (proč to můžeme?). Je $L_{\mathbf{x}}(\mathbf{x}, \lambda) = \mathbf{a}^T - \lambda \mathbf{x}^T$, stacionární body funkce L tedy splňují soustavu

$$\mathbf{a} = \lambda \mathbf{x}, \quad \mathbf{x}^T \mathbf{x} = 1.$$

Rozlišíme dva případy:

- Je-li $\lambda \neq 0$, z první rovnice máme $\mathbf{x} = \mathbf{a}/\lambda$, což dosazeno do druhé dá $\mathbf{a}^T \mathbf{a}/\lambda^2 = 1$, tedy $\lambda = \pm \|\mathbf{a}\|$, tedy kandidáti na lokální extrémy jsou $\mathbf{x} = \pm \mathbf{a}/\|\mathbf{a}\|$.
- $\lambda = 0$ může nastat jen když $\mathbf{a} = \mathbf{0}$, tedy účelová funkce je konstantní. Pak jsou kandidáti na lokální extrémy všechna \mathbf{x} splňující $\mathbf{x}^T \mathbf{x} = 1$, tedy celá sféra X . ♦

Příklad 11.16. Vraťme se k úloze (7.1), kde minimalizujeme kvadratickou formu $\mathbf{x}^T \mathbf{A}\mathbf{x}$ na sféře $\mathbf{x}^T \mathbf{x} = 1$. 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 $L_{\mathbf{x}}(\mathbf{x}, \lambda) = \mathbf{0}$ 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 [maximu] a ostatní vlastní čísla neodpovídají (globálním) extrémům úlohy. ♦

Předchozí příklady a také Příklady 11.4 a 11.3 vyžadují od studenta nejen znalost metody Lagrangeových multiplikátorů, ale i jistou zručnost v manipulaci s maticovými výrazy. Cvičte tuto zručnost ve Cvičeních 11.16–11.18!

11.2.3 Podmínky druhého řádu

Věta 11.5 udává podmínky prvního řádu na extrémů 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 . Uved’me nyní bez důkazu podmínku druhého řádu⁵. Hlavní rozdíl oproti Větě 11.3 je v tom, že místo druhých derivací funkce f v bodě \mathbf{x} v nich vystupují druhé derivace funkce L podle \mathbf{x} v bodě $(\mathbf{x}, \boldsymbol{\lambda})$, tj.

$$\frac{\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}^2} = f''(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i''(\mathbf{x}). \quad (11.24)$$

Věta 11.6. *Necht’ $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ jsou dvakrát diferencovatelné v bodě $\mathbf{x} \in \mathbb{R}^n$. Necht’ \mathbf{x} je regulární bod zobrazení \mathbf{g} .*

- *Jestliže \mathbf{x} je lokální minimum [maximum] funkce f vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, pak existuje $\boldsymbol{\lambda}$ tak že $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$ a Hessova matice (11.24) je pozitivně [negativně] semidefinitní na podprostoru $\text{null } \mathbf{g}'(\mathbf{x})$.*
- *Jestliže existuje $\boldsymbol{\lambda}$ tak že $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$ a Hessova matice (11.24) je pozitivně [negativně] definitní na podprostoru $\text{null } \mathbf{g}'(\mathbf{x})$, pak \mathbf{x} je ostré lokální minimum [maximum] funkce f vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.*

Příklad 11.17. Hledejme strany kvádrů s jednotkovým objemem 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

$$\frac{\partial^2 L(x, y, z, \lambda)}{\partial (x, y, z)^2} = \begin{bmatrix} 0 & 1 - \lambda z & 1 - \lambda y \\ 1 - \lambda z & 0 & 1 - \lambda x \\ 1 - \lambda y & 1 - \lambda x & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}. \quad (11.25)$$

Máme ukázat, že tato matice je pozitivně definitní na nulovém prostoru Jacobiho matice

$$\mathbf{g}'(x, y, z) = [-yz \quad -xz \quad -xy] = [-1 \quad -1 \quad -1]. \quad \blacklozenge$$

Dále pokračujeme podobně jako v Příkladu 11.5.

⁵Zdůrazněme, že druh lokálního extrému *nelze* zjistit podle definitnosti Hessovy matice $L''(\mathbf{x}, \boldsymbol{\lambda})$, tedy je chybou použít Větu 10.4 na funkci L (z Věty 11.6 nic takového neplatí).

11.3 Cvičení

Následující úlohy vyřešte nejdříve převodem na hledání volných extrémů dle Příkladu 11.1 (proved'te jen je-li to možné; někdy je k tomu třeba jednoduchá úvaha) 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 extrémů nějakou snadnou úvahou, udělejte to).

11.1. 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) $f(x, y) = e^{xy}$

11.2. 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) $(\star) f(x, y, z) = x^3 + 2xyz - z^3$

11.3. Najděte lokální extrémy funkce

- a) $f(x, y, z) = x + yz$ za podmíněk $x^2 + y^2 + z^2 = 1$ a $z^2 = x^2 + y^2$
- b) $f(x, y, z) = xyz$ za podmíněk $x^2 + y^2 + z^2 = 1$ a $xy + yz + zx = 1$

11.4. Najděte bod nejbližší počátku na křivce

- a) $x + y = 1$
- b) $x + 2y = 5$
- c) $x^2y = 1$
- d) $x^2 + 2y^2 = 1$

11.5. 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) půllitr (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 vhodnou numerickou metodu na řešení vzniklé soustavy rovnic.

- 11.6. Rozložte dané kladné reálné číslo na součin n kladných reálných čísel tak, aby jejich součet byl co nejmenší.
- 11.7. Najděte vzdálenost množiny $\{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$ od kružnice s poloměrem 1 a středem v bodě $(2, 0)$.
- 11.8. 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.
- 11.9. 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}^* .
- 11.10. Do elipsy o daných délkách os vepište obdélník s maximálním obsahem. Předpokládejte přitom, že strany obdélníku jsou rovnoběžné s osami elipsy.
- 11.11. *Fermatův princip* v paprskové optice říká, že cesta mezi libovolnými dvěma body na paprsku má takový tvar, aby ji světlo proběhlo za čas kratší než jí blízké dráhy. Později se zjistilo, že správným kritériem není *nejkratší* ale *extrémní* čas. Tedy skutečná dráha paprsku musí mít čas větší nebo menší než jí blízké dráhy. Z tohoto principu odvoďte:

- a) Zákon odrazu od zrcadla: úhel dopadu se rovná úhlu odrazu.
 b) Snellův zákon lomu: na rozhraní dvou prostředí se světlo lomí tak, že

$$\frac{c_1}{c_2} = \frac{\sin \alpha_1}{\sin \alpha_2},$$

kde α_i je úhel paprsku od normály rozhraní a c_i je rychlost světla v prostředí i .

Odvození udělejte:

- (i) Pro rovinné zrcadlo a rovinné rozhraní (což vede na minimalizaci bez omezení).
 (ii) Pro zrcadlo a rozhraní tvaru obecné plochy s rovnicí $g(\mathbf{x}) = 0$. Dokážete najít situaci, kdy skutečná dráha paprsku má čas *větší* než jí blízké dráhy?
- 11.12. Rozdělení pravděpodobnosti diskrétní náhodné proměnné je funkce $p: \{1, \dots, n\} \rightarrow \mathbb{R}_+$ (tj. soubor nezáporných čísel $p(1), \dots, p(n)$) splňující $\sum_{x=1}^n p(x) = 1$.
- a) *Entropie* náhodné proměnné s rozdělením p je rovna $-\sum_{x=1}^n p(x) \log p(x)$, kde \log je přirozený logaritmus. Najděte rozdělení s maximální entropií.
 b) Dokažte *Gibbsovu nerovnost* (též zvanou *informační nerovnost*): pro každé dvě rozdělení p, q platí

$$\sum_{x=1}^n p(x) \log q(x) \leq \sum_{x=1}^n p(x) \log p(x),$$

přičemž rovnost nastává jen tehdy, když $p = q$.

- 11.13. (★) 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?
- 11.14. (★) 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.

11.15. (★) 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 $x, y, z, u, v, w > 0$.

11.16. Minimalizujte $\mathbf{x}^T \mathbf{x}$ za podmínky $\mathbf{a}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?

11.17. Minimalizujte $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\mathbf{b}^T \mathbf{x} = 1$, kde \mathbf{A} je symetrická pozitivně definitní.

11.18. 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. Najděte vzorec pro optimální \mathbf{x} .

11.19. Řešte Cvičení 11.18, kde ale matice \mathbf{C} a \mathbf{A} mohou být libovolné. Nemusíte najít vzorec pro optimální \mathbf{x} , ale napište lineární soustavu, jejímž řešením je stacionární bod Lagrangeovy funkce (jako v Příkladu 11.4).

11.20. (★) 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.

11.21. (★) 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.

11.22. (★) Minimalizujte $\mathbf{a}^T (\mathbf{x} - \mathbf{b})$ za podmínky $\mathbf{x}^T \mathbf{C} \mathbf{x} = 1$, kde \mathbf{C} je pozitivně definitní.

11.23. (★) Pro jaké matice \mathbf{A} a vektory \mathbf{b} platí $\max\{\|\mathbf{A} \mathbf{x}\| \mid \mathbf{b}^T \mathbf{x} = 0\} = 0$?

11.24. (★) 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 11.4).

11.25. Necht' $g: \mathbb{R}^n \rightarrow \mathbb{R}$ je všude diferencovatelná a všude má nenulový gradient. Necht' $\mathbf{0} \neq \mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Hledáme vzdálenost křivého povrchu $g(\mathbf{x}) = 0$ od nadroviny $\mathbf{a}^T \mathbf{x} = b$. Dokažte, že tečný prostor k nadpovrchu v bodě nejbližší nadrovině bude rovnoběžný s nadrovinou.

11.26. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná v bodě $\mathbf{x}^* \in \mathbb{R}^n$. Ukažte, že graf Taylorova polynomu $T_{\mathbf{x}^*}^1(\mathbf{x}) = f(\mathbf{x}^*) + f'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$ prvního stupně (tj. afinní aproximace) funkce f v bodě \mathbf{x}^* je tečný podprostor ke grafu funkce f (zopakujte definici grafu funkce ze zač. §8!) v bodě \mathbf{x}^* , přičemž podmínky regularity jsou splněny.

Nápověda a řešení

11.1.d) Můžeme použít stejnou parametrizaci kružnice jako v Příkladě 11.1 a najít (s trochou štěstí) čtyři lokální extrémy funkce $\cos^2 \alpha \sin \alpha$. Můžeme ale užít i jiný způsob. Do účelové funkce dosadíme $x^2 = 1 - y^2$ a hledáme lokální extrémy funkce $(1 - y^2)y$. Ale pozor: hledáme je ne na množině \mathbb{R} , ale na intervalu $[-1, 1]$. Na tomto intervalu má funkce dva lokální extrémy $y = \pm \frac{1}{\sqrt{3}}$ uvnitř intervalu a dva extrémy ± 1 v krajních bodech intervalu. To dává celkem šest lokálních extrémů (x, y) původní úlohy.

11.1.g) 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.

11.4.c) Jsou dva body nejbližší počátku: $x = \pm 2^{1/6}$, $y = 2^{-1/3}$.

11.6. Formulace: Minimalizujeme $x_1 + \dots + x_n$ za podmínek $x_1 \dots x_n = a$ a $x_1, \dots, x_n > 0$, kde $a > 0$ je dané číslo.

11.7. Ukážeme jen řešení pomocí převodu na volné extrémy: po eliminaci proměnné y minimalizujeme $\|(x, x^2) - (2, 0)\|_2^2 = (x - 2)^2 + x^4$. Stacionární podmínka je $2x^3 + x - 2 = 0$. To je kubická rovnice. Pokud neznáme vzorce na řešení kubických rovnic (tzv. Cardanovy vzorce), pomůže Newtonova

metoda: její iterace je (odvoďte!) $x_{k+1} \leftarrow x_k - \frac{2x_k^3 + x_k - 2}{6x_k^2 + 1} = \frac{4x_k^3 + 2}{6x_k^2 + 1}$, počáteční odhad $x_0 = 1$.

Po pár iteracích je $x \approx 0.835122348481367$, tedy $d \approx 1.357699386102247$.

11.11.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 pílí úhel mezi nimi (nakreslete si).

11.12.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.)

$$11.17. \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}/(\mathbf{b}^T \mathbf{A}^{-1}\mathbf{b})$$

$$11.18. \mathbf{x} = \mathbf{C}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{C}^{-1}\mathbf{A}^T)^{-1}\mathbf{b}.$$

11.25. Minimalizujeme $\|\mathbf{x} - \mathbf{y}\|^2$ za podmínek $g(\mathbf{x}) = 0$ a $\mathbf{a}^T \mathbf{y} = b$. Dokazované tvrzení vyplývá z podmínek prvního řádu (pomocí Lagrangeových multiplikátorů).

11.26. Toto cvičení se týká §11.2.1. Graf funkce f je množina $X = \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid g(\mathbf{x}, y) = 0\}$, kde $g: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ je $g(\mathbf{x}, y) = f(\mathbf{x}) - y$. Jacobiho matice $g'(\mathbf{x}, y) = [f'(\mathbf{x}) \quad -1]$ je vždy nenulový řádkový vektor, tedy každý bod $(\mathbf{x}, f(\mathbf{x})) \in X$ je regulární bod funkce g .

Graf funkce $T_{\mathbf{x}^*}^1$ je afinní podprostor $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}^*) + f'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = y\}$. Posuneme-li ho do počátku, stane se z něj (lineární) podprostor $\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f'(\mathbf{x}^*)\mathbf{x} = y\}$. Máte ukázat, že tento podprostor je roven tečnému podprostoru (bez posunutí) k množině X v bodě $(\mathbf{x}^*, f(\mathbf{x}^*)) \in X$, tedy $\text{null}[f'(\mathbf{x}^*) \quad -1]$.

Část III

Lineární programování

Kapitola 12

Lineární programování

Úloha **lineárního programování**¹ (LP, také zvané lineární optimalizace) je minimalizace nebo maximalizace lineární funkce za omezujících podmínek ve tvaru lineárních rovnic a nerovnic. Zde **lineární rovnici** rozumíme relaci

$$a_1x_1 + \cdots + a_nx_n = b,$$

neboli krátce $\mathbf{a}^T \mathbf{x} = b$. **Lineární nerovnici** rozumíme jednu z relací

$$a_1x_1 + \cdots + a_nx_n \leq b, \quad a_1x_1 + \cdots + a_nx_n \geq b,$$

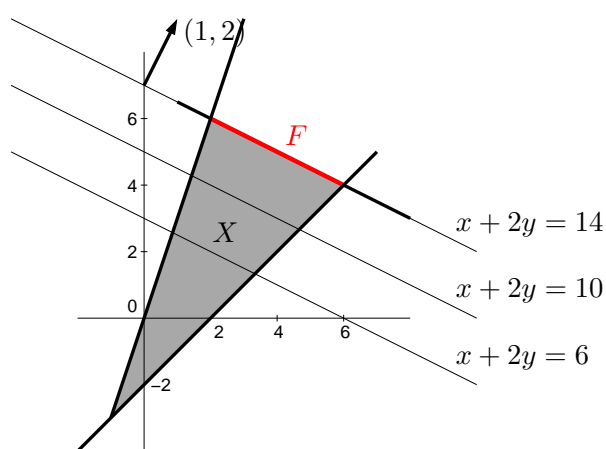
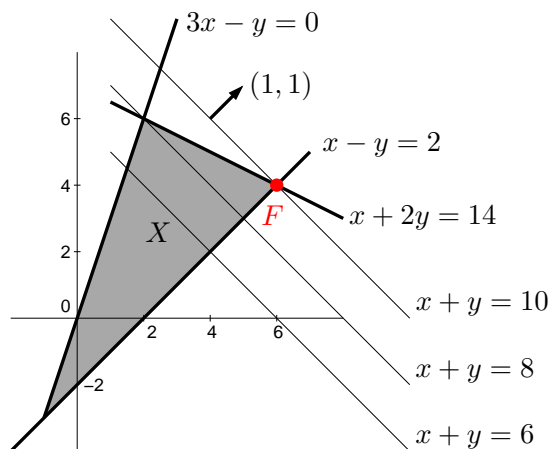
neboli $\mathbf{a}^T \mathbf{x} \leq b$ či $\mathbf{a}^T \mathbf{x} \geq b$. Lineární program je tedy úloha (1.10) ve které je funkce f lineární (tj. tvaru (3.11)) a funkce g_i, h_i jsou afinní (tj. tvaru (3.37)).

Lineární programy s jednou nebo dvěma proměnnými lze řešit graficky.

Příklad 12.1. Příkladem lineárního programu je optimalizační úloha

$$\begin{aligned} \max \quad & x + y \\ \text{za podmínek} \quad & x + 2y \leq 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \end{aligned} \tag{12.1}$$

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 . ♦

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.

12.1 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):

$$\begin{aligned} \min \quad & c_1 x_1 + \cdots + c_n x_n \\ \text{za podmíněk} \quad & a_{i1} x_1 + \cdots + a_{in} x_n \geq b_i, \quad i = 1, \dots, m \end{aligned} \quad (12.2a)$$

To se pohodlněji napíše v maticovém tvaru

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \mathbf{x} \geq \mathbf{b} \}, \quad (12.2b)$$

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 pro každé $i = 1, \dots, m$ není i -tá složka vektoru $\mathbf{A} \mathbf{x}$ menší než i -tá složka vektoru \mathbf{b} . Jakýkoliv lineární program snadno převedeme na tvar (12.2) těmito úpravami:

- Maximalizaci funkce $\mathbf{c}^T \mathbf{x}$ nahradíme minimalizací funkce $-\mathbf{c}^T \mathbf{x}$.
- Nerovnost $\mathbf{a}^T \mathbf{x} \leq b$ nahradíme nerovností $-\mathbf{a}^T \mathbf{x} \geq -b$.
- 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

$$\begin{aligned} \min \quad & c_1 x_1 + \cdots + c_n x_n \\ \text{za podmíněk} \quad & a_{i1} x_1 + \cdots + a_{in} x_n = b_i, \quad i = 1, \dots, m \\ & x_j \geq 0, \quad j = 1, \dots, n \end{aligned} \quad (12.3a)$$

neboli

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A} \mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}. \quad (12.3b)$$

Tvar (12.2) lze převést na tvar (12.3) dvěma úpravami, ve kterých zavedeme nové proměnné:

²Tomuto tvaru se někdy říká *standardní*. Názvosloví bohužel není jednotné, názvy jako ‘standardní tvar’, ‘základní tvar’ či ‘kanonický tvar’ tedy mohou znamenat v různých knihách různé věci.

- Nerovnost $\mathbf{a}^T \mathbf{x} \geq b$ nahradíme dvěma omezeními $\mathbf{a}^T \mathbf{x} - u = b$, $u \geq 0$. Pomocné proměnné u se říká **slacková proměnná**³. Podobně bychom převedli nerovnost $\mathbf{a}^T \mathbf{x} \leq b$ 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 12.2. Úlohu (12.1) 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 (12.3), kde všechny proměnné jsou nezáporné. Dosadíme $x = x_+ - x_-$ a $y = y_+ - y_-$, kde $x_+, x_-, y_+, y_- \geq 0$. Výsledná úloha je

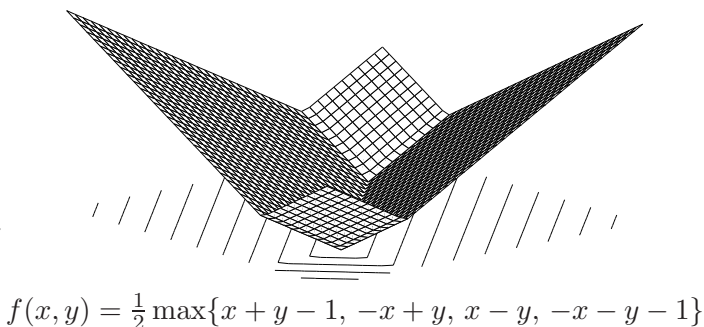
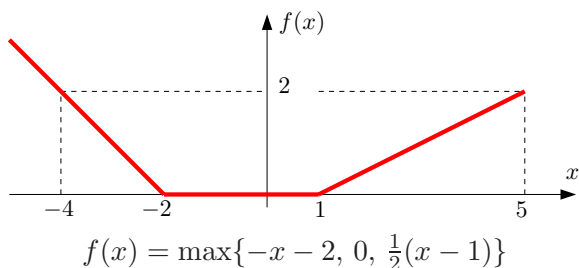
$$\begin{aligned} \min \quad & -x_+ + x_- - y_+ + y_- \\ \text{za podmínek} \quad & x_+ - x_- + 2y_+ - 2y_- + u = 14 \\ & 3x_+ - 3x_- - y_+ + y_- - v = 0 \\ & x_+ - x_- - y_+ + y_- + w = 2 \\ & x_+, x_-, y_+, y_-, u, v, w \geq 0 \end{aligned} \quad \blacklozen$$

12.1.1 Po částech afinní funkce

Mějme funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ danou vzorcem

$$f(\mathbf{x}) = \max_{i=1}^k (\mathbf{c}_i^T \mathbf{x} + d_i), \quad (12.4)$$

kde $\mathbf{c}_1, \dots, \mathbf{c}_k \in \mathbb{R}^n$ a $d_1, \dots, d_k \in \mathbb{R}$ 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 přiřazená ke zdi. Termín *slack variable* nemá ustálený český ekvivalent, někdy se překládá jako *skluzová proměnná*.

⁴Přesněji v *lineárním čase*, ve smyslu teorie algoritmů.

Minimalizujme funkci (12.4) 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.

Obecně, pro každou množinu X a funkci $f: X \rightarrow \mathbb{R}$ platí (pokud minimum existuje)

$$\min\{f(x) \mid x \in X\} = \min\{y \mid (x, y) \in X \times \mathbb{R}, f(x) \leq y\}. \quad (12.5)$$

To proto, že každé optimální řešení (x, y) pravé úlohy splňuje $f(x) = y$. Kdyby totiž bylo $f(x) < y$, pak by řešení (x, y) bylo přípustné ale ne optimální, protože bychom mohli y zmenšit bez porušení omezení a zlepšit tak účelovou funkci.

Tedy už naši úlohu snadno převedeme na LP:

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \geq \mathbf{b}\} = \min\{y \mid (\mathbf{x}, y) \in \mathbb{R}^{n+1}, f(\mathbf{x}) \leq y, \mathbf{Ax} \geq \mathbf{b}\} \quad (12.6a)$$

$$= \min\{y \mid (\mathbf{x}, y) \in \mathbb{R}^{n+1}, \max_i(\mathbf{c}_i^T \mathbf{x} + d_i) \leq y, \mathbf{Ax} \geq \mathbf{b}\} \quad (12.6b)$$

$$= \min\{y \mid (\mathbf{x}, y) \in \mathbb{R}^{n+1}, (\forall i)(\mathbf{c}_i^T \mathbf{x} + d_i \leq y), \mathbf{Ax} \geq \mathbf{b}\} \quad (12.6c)$$

kde (12.6c) je již úloha LP. Rovnost (12.6c) platí proto, že pro libovolná čísla a_i, b zřejmě je

$$\max_i a_i \leq b \iff (\forall i)(a_i \leq b). \quad (12.7)$$

Příklad 12.3. Úloha

$$\begin{aligned} & \min \max\{3x + 4y + 1, 2x - 3y\} \\ \text{za podm. } & \begin{aligned} x + 2y &\leq 14 \\ 3x - y &\geq 0 \\ x - y &\leq 2 \end{aligned} \end{aligned}$$

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

$$\begin{aligned} & \min z \\ \text{za podm. } & \begin{aligned} 3x + 4y - z &\leq -1 \\ 2x - 3y - z &\leq 0 \\ x + 2y &\leq 14 \\ 3x - y &\geq 0 \\ x - y &\leq 2 \end{aligned} \end{aligned} \quad \blacklozenge$$

Tento převod lze užít i pro funkce obsahující absolutní hodnoty, neboť $|x| = \max\{-x, x\}$, a na různé jiné případy (trénujte to ve Cvičení 12.4!). Bud'te 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.

12.2 Jednoduché úlohy LP

Některé úlohy LP lze řešit úvahou. Je dobré se v tom pocvičit, protože vám to dá intuici potřebnou i pro obecnější optimalizační úlohy. Více viz Cvičení 12.3.

Příklad 12.4. Pro daná čísla $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^n$ řešíme úlohu

$$\begin{aligned} \max \quad & c_1 x_1 + \dots + c_n x_n \\ \text{za podmíněk} \quad & x_1 + \dots + x_n = 1 \\ & x_1, \dots, x_n \geq 0 \end{aligned}$$

neboli

$$\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}.$$

Přípustná řešení je možno vidět jako rozdělení pravděpodobnosti diskrétní náhodné veličiny s n stavy. Není těžké uhodnout (zamyslete se!), že optimum úlohy se nabývá právě tehdy, když $x_i = 0$ pro všechna $i \notin \operatorname{argmax}_i c_i$ (a zbývající proměnné jsou libovolně přípusté). Jestliže množina $\operatorname{argmax}_i c_i$ má pouze jeden prvek i^* , pak $x_i = 0$ pro všechna $i \neq i^*$ a $x_{i^*} = 1$. V obou případech je optimální hodnota rovna číslu $c_{\max} = \max_i c_i$, tj. největšímu z čísel c_1, \dots, c_n .

Jak dokážeme, že uhodnuté řešení je skutečně optimální? Tak, že ukážeme, že každé jiné přípustné řešení bychom mohli pozměnit tak, že by zůstalo přípustné a účelová funkce by se zlepšila. Kdyby např. bylo $x_1 > 0$ a $c_1 < c_{\max}$, pak bychom mohli x_1 o kousek zmenšit a tento kousek přidat k x_i pro nějaké i splňující $c_i = c_{\max}$, což by zlepšilo (zvýšilo) účelovou funkci.

Uvedeme ještě jiný důkaz, že optimální hodnota úlohy je c_{\max} . Hodnota kritéria v libovolném přípustném řešení je *dolní mez* na optimální hodnotu úlohy. Pro naše uhodnuté přípustné řešení máme tedy dolní mez c_{\max} . Ale pro každé přípustné řešení x_1, \dots, x_n je

$$c_1 x_1 + \dots + c_n x_n \leq c_{\max} x_1 + \dots + c_{\max} x_n = c_{\max} (x_1 + \dots + x_n) = c_{\max}, \quad (12.8)$$

tedy c_{\max} je zároveň *horní mez* na optimální hodnotu úlohy. Protože obě meze splývají, optimální hodnota musí být c_{\max} . ♦

Příklad 12.5. Pozměňme předcházející úlohu: jsou dány $\mathbf{c} \in \mathbb{R}^n$ a $k \in \{1, \dots, n\}$ a řešíme

$$\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{1}^T \mathbf{x} = k, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}.$$

Bez ztráty obecnosti předpokládejme, že čísla c_i jsou sestupně seřazena, $c_1 \geq \dots \geq c_n$.

Opět není těžké uhodnout, že optimum úlohy se nabývá právě tehdy, když $x_i = 0$ pro všechna i splňující $c_i < c_k$ (a ostatní proměnné jsou libovolně přípustné). Tomu odpovídá optimální hodnota $c_1 + \dots + c_k$, tedy součet k největších čísel c_i . Speciálně, kdyby $c_{k+1} < c_k$, pak by optimum nastalo právě pro $x_1 = \dots = x_k = 1$ a $x_{k+1} = \dots = x_n = 0$.

Optimalitu tohoto řešení dokážeme podobným argumentem jako minule: kdyby pro nějaké i bylo $x_i > 0$ a $c_i < c_k$, pak by nutně (kvůli podmínce $x_1 + \dots + x_n = k$) bylo $x_{i^*} < 1$ pro nějaké i^* splňující $c_{i^*} \geq c_k$ a mohli bychom tedy x_i o kousek zmenšit a x_{i^*} o stejný kousek zvětšit, což by zvýšilo hodnotu kritéria.

I když důkaz z předchozího odstavce dostačuje, odvodíme i horní mez analogicky k (12.8): pro každé přípustné \mathbf{x} platí (promyslete, proč každá rovnost a nerovnost platí!)

$$\sum_{i=1}^n c_i x_i \leq \sum_{i=1}^k c_i x_i + c_k \sum_{i=k+1}^n x_i = \sum_{i=1}^k c_i x_i + c_k \sum_{i=1}^k (1 - x_i) = \sum_{i=1}^k \underbrace{(c_i x_i + c_k (1 - x_i))}_{\leq c_i} \leq \sum_{i=1}^k c_i. \quad \blacklozenge$$

12.3 Typické aplikace LP

Ve zbytku této kapitoly uvedeme některá z mnoha použití lineárního programování.

12.3.1 Optimální výrobní program

Z m druhů surovin vyrábíme n druhů výrobků.

- a_{ij} = množství suroviny druhu i potřebné na výrobu výrobku druhu j
- b_i = množství suroviny druhu i , které máme k dispozici
- c_j = zisk z vyrobení jednoho výrobku druhu j
- x_j = počet vyrobených výrobků druhu j

Chceme zjistit, kolik jakých výrobků máme vyrobit, aby zisk byl největší. Tuto úlohu lze formalizovat lineárním programem

$$\max \sum_{j=1}^n c_j x_j \quad (12.9a)$$

$$\text{za podm. } \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \dots, m \quad (12.9b)$$

$$x_j \geq 0, \quad j = 1, \dots, n \quad (12.9c)$$

což jde kratěji napsat v maticové formě jako $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

Příkladem úlohy na optimální výrobní program je Příklad 1.10 z Kapitoly 1 (podívejte se na něj!).

12.3.2 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

$$\min \sum_{j=1}^n c_j x_j \quad (12.10a)$$

$$\text{za podm. } \sum_{j=1}^n a_{ij} x_j \geq b_i, \quad i = 1, \dots, m \quad (12.10b)$$

$$x_j \geq 0, \quad j = 1, \dots, n \quad (12.10c)$$

neboli $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

Příklad 12.6. Jste kuchařka v menze a chcete uvařit pro studenty co nejlevnější oběd, ve kterém ovšem kvůli předpisům musí být dané minimální množství živin (cukrů, bílkovin a vitamínů). Oběd 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

$$\begin{aligned} \min \quad & 25b + 50m + 80z \\ \text{za podmíněk} \quad & 2b + m + z \geq 8 \\ & 2b + 6m + z \geq 16 \\ & b + 3m + 6z \geq 8 \\ & b, m, z \geq 0 \end{aligned}$$

Optimální řešení je $b = 3.2$, $m = 1.6$, $z = 0$ s hodnotou 160. ♦

12.3.3 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

$$\min \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \tag{12.11a}$$

$$\text{za podm.} \quad \sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m \tag{12.11b}$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n \tag{12.11c}$$

$$x_{ij} \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n \tag{12.11d}$$

Ten jde také napsat jako (rozmyslete!)

$$\min \{ \langle \mathbf{C}, \mathbf{X} \rangle \mid \mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{X}\mathbf{1} = \mathbf{a}, \mathbf{X}^T\mathbf{1} = \mathbf{b}, \mathbf{X} \geq \mathbf{0} \}, \tag{12.12}$$

kde $\mathbf{a} = (a_1, \dots, a_m)$, $\mathbf{b} = (b_1, \dots, b_n)$, $\mathbf{C} = [c_{ij}]$, $\langle \mathbf{C}, \mathbf{X} \rangle$ značí skalární součin matic (2.7) 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 nepřijatelná. To snadno vidíme: součet levých stran rovnic (12.11b) je roven součtu levých stran rovnic (12.11c). 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 (12.11b) muselo změnit na $\sum_{j=1}^n x_{ij} \leq a_i$ (promyslete!).

12.3.4 Distribuční úloha

Máme m strojů a n druhů výrobků.

- a_i = počet hodin, který je k dispozici na stroji i
- b_j = požadované množství výrobku druhu j
- c_{ij} = cena jedné hodiny práce stroje i na výrobku typu j
- k_{ij} = hodinový výkon stroje i při výrobě výrobku druhu j
- x_{ij} = počet hodin, po který bude stroj i vyrábět výrobek druhu j

Pro každý ze strojů máme určit, kolik výrobků se na něm bude vyrábět. Řešení:

$$\min \left\{ \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \mid \sum_{j=1}^n x_{ij} \leq a_i, \sum_{i=1}^m k_{ij} x_{ij} = b_j, x_{ij} \geq 0 \right\}. \quad (12.13)$$

12.4 Přeuročené lineární soustavy

12.4.1 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**⁵, 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 $\mathbf{y} = -\mathbf{x}$ a tedy

$$0 = \|\mathbf{0}\| = \|\mathbf{x} - \mathbf{x}\| \leq \|\mathbf{x}\| + \|-\mathbf{x}\| = 2\|\mathbf{x}\|,$$

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**

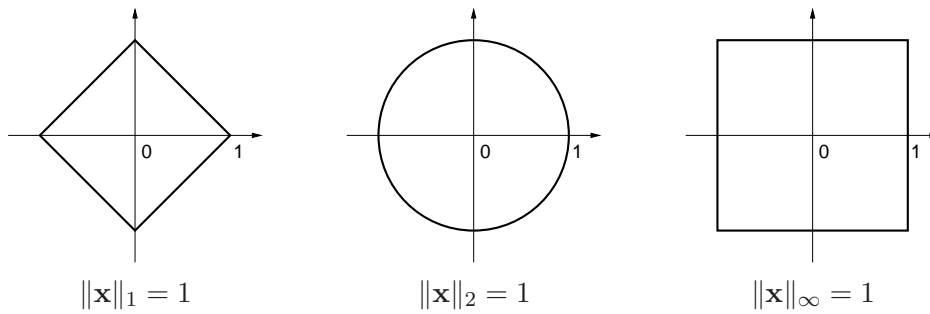
$$\|\mathbf{x}\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p}.$$

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*.
- $\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \max\{|x_1|, \dots, |x_n|\}$ (spočítejte limitu ve Cvičení 12.13!). Někdy se jí říká *Čebyševova norma* nebo *max-norma*.

⁵Symbolem $\|\cdot\|$ jsme dosud značili eukleidovskou normu. Od teď až do konce skript jím ale budeme značit obecnou vektorovou normu a eukleidovskou normu budeme značit symbolem $\|\cdot\|_2$.

Jednotkové sféry těchto norm v \mathbb{R}^2 vypadají takto:



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 $\|\mathbf{Ax}\|$ norma.

12.4.2 Přibližné řešení lineárních soustav v 1-normě a ∞ -normě

Mějme přeурčenu lineární soustavu $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{0} \neq \mathbf{b} \in \mathbb{R}^m$. Nalezení jejího přibližného řešení formulujeme jako úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_p. \quad (12.14)$$

Uvažujme tři případy:

- Pro $p = \infty$ hledáme takové \mathbf{x} , které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_\infty = \max_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (12.15)$$

tedy minimalizuje maximální residuum. Toto řešení je známé pod názvem *minimaxní* nebo *Čebyševovo*. Podle §12.1.1 je tato úloha ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & -z \leq \mathbf{a}_i^T \mathbf{x} - b_i \leq z, \quad i = 1, \dots, m \end{aligned}$$

což lze napsat kratěji v maticové formě

$$\min\{z \mid \mathbf{x} \in \mathbb{R}^n, z \in \mathbb{R}, -z\mathbf{1} \leq \mathbf{Ax} - \mathbf{b} \leq z\mathbf{1}\}. \quad (12.16)$$

- 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

$$\|\mathbf{Ax} - \mathbf{b}\|_1 = \sum_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (12.17)$$

kde $\mathbf{a}_1^T, \dots, \mathbf{a}_m^T$ jsou řádky matice \mathbf{A} . Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & z_1 + \dots + z_m \\ \text{za podm.} \quad & -z_i \leq \mathbf{a}_i^T \mathbf{x} - b_i \leq z_i, \quad i = 1, \dots, m \end{aligned}$$

neboli

$$\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^m, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}\}. \quad (12.18)$$

12.4.3 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

$$\sum_{i=1}^m (y_i - f(x_i, \boldsymbol{\theta}))^2 = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_2, \quad (12.19)$$

kde $\mathbf{y} = (y_1, \dots, y_m)$ a prvky matice \mathbf{A} jsou $a_{ij} = \varphi_j(x_i)$. Tedy řešíme úlohu (12.14) pro $p = 2$. Můžeme ale použít i jiné normy než eukleidovskou. Pro $p = 1$ minimalizujeme

$$\sum_{i=1}^m |y_i - f(x_i, \boldsymbol{\theta})| = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_1 \quad (12.20)$$

a pro $p = \infty$ minimalizujeme

$$\max_{i=1}^m |y_i - f(x_i, \boldsymbol{\theta})| = \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_\infty. \quad (12.21)$$

Dále ukážeme, k čemu to může být dobré.

Regrese ve smyslu ∞ -normy je vhodná např. při aproximaci funkcí.

Příklad 12.7. Na počítači bez matematického koprocesoru 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 $i = 1, \dots, m$. Koefficienty polynomu hledáme minimalizací Čebyševova kritéria (12.21), neboť to nám dá záruku, že chyba aproximace nikde nepřesáhne hodnotu, která je nejmenší možná pro daný stupeň polynomu. \blacklozenge

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

$$\|\mathbf{y} - \mathbf{1}\theta\|_p = \|(y_1 - \theta, \dots, y_m - \theta)\|_p. \quad (12.22)$$

- 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 Příklad 1.13.

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 12.8. 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

$$\frac{1}{2} \left(\min_{i=1}^m y_i + \max_{i=1}^m y_i \right) = 1.51, \quad \frac{1}{m} \sum_{i=1}^m y_i = 1.27, \quad \text{median}_{i=1}^m y_i = 1.03.$$

Je zjevné, že medián je neovlivněn vychýleným bodem, zatímco ostatní odhady ano. ◆

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.

12.5 Celočíslné lineární programování, LP relaxace

Celočíslné 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

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{A} \mathbf{x} \geq \mathbf{b} \}. \quad (12.23)$$

Množina přípustných řešení této úlohy obsahuje konečný (avšak obvykle obrovský) počet izolovaných bodů, jedná se o typickou úlohu *kombinatorické optimalizace* (viz §1.3). Zatímco vyřešit obyčejný lineární program je relativně snadné (LP je řešitelné v polynomiálním čase), vyřešit úlohu ILP je obecně velmi obtížné (ILP je tzv. *NP-těžké*).

Mohli bychom zkusit úlohu (12.23) jednoduše nahradit lineárním programem

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in [0, 1]^n, \mathbf{A} \mathbf{x} \geq \mathbf{b} \}. \quad (12.24)$$

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 (12.24) se říká **LP relaxace** (angl. *relaxation* znamená uvolnění, myšleno jako ‘uvolnění’ omezení) úlohy (12.23). 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 (12.24) nějak šikovně zaokrouhlíme na celočíselný vektor $\mathbf{x}' \in \{0, 1\}^n$, dostaneme řešení (nebo aspoň přibližné řešení) úlohy (12.23). 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 (12.23) je přípustná (tedy zda existuje alespoň jedno $\mathbf{x} \in \{0, 1\}^n$ splňující $\mathbf{Ax} \geq \mathbf{b}$), je tzv. *NP-úplný problém* (zhruba řečeno, je to stejně těžké jako najít optimální řešení úlohy (12.23)). I kdybychom 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 (12.23).

Obecně, jediný užitečný vztah mezi optimálním řešením původní a relaxované úlohy je, že optimální hodnota relaxované úlohy není větší než optimální hodnota původní úlohy:

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in [0, 1]^n, \mathbf{Ax} \geq \mathbf{b}\} \leq \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{Ax} \geq \mathbf{b}\}. \quad (12.25)$$

To proto, že v relaxované úloze minimalizujeme přes *větší* množinu než v původní úloze:

$$\{\mathbf{x} \in [0, 1]^n \mid \mathbf{Ax} \geq \mathbf{b}\} \supseteq \{\mathbf{x} \in \{0, 1\}^n \mid \mathbf{Ax} \geq \mathbf{b}\}.$$

Přesněji to lze uvidět takto: Necht' \mathbf{x}^* je optimální řešení nerelaxované úlohy (12.23) a necht' \mathbf{x} je optimální řešení relaxované úlohy (12.24). Protože \mathbf{x}^* je přípustné (i když ne nutně optimální) pro relaxovanou úlohu, musí být $\mathbf{c}^T \mathbf{x} \leq \mathbf{c}^T \mathbf{x}^*$. Rozdílu mezi pravou a levou stranou nerovnosti (12.25) se říká *integrality gap* (tedy *mezera celočíselnosti*).

Pro některé konkrétní kombinatorické problémy však je situace příznivější, např. LP relaxace řeší původní úlohu vždy přesně (tedy nerovnost (12.25) platí s rovností), nebo dovoluje sestrojít její přibližné řešení se zárukou přesnosti (dále uvedeme příklady). I když toto štěstí nemáme, dolní mez je často užitečná při hledání přesného nebo přibližného řešení úlohy nějakým jiným způsobem (např. metodou *větví a mezí* nebo metodou *sečných nadrovin*).

12.5.1 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 minimalizuje 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

$$\min \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \quad (12.26a)$$

$$\text{za podm. } \sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n \quad (12.26b)$$

$$\sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n \quad (12.26c)$$

$$x_{ij} \in \{0, 1\}, \quad i, j = 1, \dots, n \quad (12.26d)$$

LP relaxaci úlohy (12.26) získáme nahrazením omezení $x_{ij} \in \{0, 1\}$ omezeními $0 \leq x_{ij} \leq 1$. Vlastně stačí jen $x_{ij} \geq 0$, neboť (12.26b) a $x_{ij} \geq 0$ implikuje $x_{ij} \leq 1$. Takže LP relaxace je speciální případ dopravního problému (12.11).

Pro tuto úlohu máme stěší, protože platí toto (důkaz neuvádíme, není krátký):

Věta 12.1. Úloha (12.26) 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 (12.26a) je lineární funkce matice \mathbf{X} (jde psát jako $\langle \mathbf{C}, \mathbf{X} \rangle$, viz §2.1.7). Přípustná řešení úlohy (12.26) jsou všechny matice, které mají v každém řádku a každém sloupci právě jednu jedničku. To jsou *permutační matice* (viz Příklad 4.3). 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*. Každá permutační matice je dvojitě stochastická, ale naopak to neplatí. Věta 12.1 říká, že minimum lineární funkce na množině permutačních matic je rovno minimu té samé lineární funkce na množině dvojitě stochastických matic. To je klasický výsledek *Birkhoffa a von Neumanna*.

12.5.2 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

$$\min \sum_{i \in V} x_i \quad (12.27a)$$

$$\text{za podm. } x_i + x_j \geq 1, \quad \{i, j\} \in E \quad (12.27b)$$

$$x_i \in \{0, 1\}, \quad i \in V \quad (12.27c)$$

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 (12.27) můžeme zapsat také jako

$$\min \{ \mathbf{1}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{A}^T \mathbf{x} \geq \mathbf{1} \}. \quad (12.28)$$

LP relaxace úlohy (12.27) se získá tak, že omezení $x_i \in \{0, 1\}$ se nahradí omezeními $0 \leq x_i \leq 1$.

Necht' x_i ($i \in V$) je optimální řešení relaxované úlohy. Zaokrouhleme toto řešení:

$$\bar{x}_i = \left\lfloor x_i + \frac{1}{2} \right\rfloor = \begin{cases} 0 & \text{když } x_i < \frac{1}{2}, \\ 1 & \text{když } x_i \geq \frac{1}{2}. \end{cases}$$

Zaokrouhlené řešení je přípustné pro úlohu (12.27): protože $x_i + x_j \geq 1$, musí být $x_i \geq \frac{1}{2}$ nebo $x_j \geq \frac{1}{2}$, tedy $\bar{x}_i = 1$ nebo $\bar{x}_j = 1$, tedy $\bar{x}_i + \bar{x}_j \geq 1$. Označme optimální hodnotu nerelaxované úlohy (12.27) jako y^* , optimální hodnotu relaxované úlohy jako $y = \sum_{i \in V} x_i$ a hodnotu zaokrouhleného řešení relaxace jako $\bar{y} = \sum_{i \in V} \bar{x}_i$.

Věta 12.2. Pro každý problém minimálního vrcholového pokrytí je $\bar{y} \leq 2y^*$.

Důkaz. Máme

$$\bar{y} = \sum_{i \in V} \bar{x}_i \leq 2 \sum_{i \in V} x_i = 2y \leq 2 \sum_{i \in V} x_i^* = 2y^*.$$

První nerovnost plyne z toho, že pro každé číslo $x \geq 0$ je $\lfloor x + \frac{1}{2} \rfloor \leq 2x$, tedy $\bar{x}_i \leq 2x_i$. Druhá nerovnost plyne z (12.25). ■

Celkově tedy máme tyto nerovnosti (rozmyslete):

$$0 \leq y \leq y^* \leq \bar{y} \leq 2y.$$

Z toho plyne $\frac{1}{2}y^* \leq y \leq y^*$, tedy optimální hodnota LP relaxace nemůže být libovolně-krát menší než optimální hodnota nerelaxované úlohy (mezera celočíselnosti nemůže být libovolná).

Příklad 12.9. Necht' $V = \{1, 2, 3\}$ a $E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$, tedy (V, E) je úplný graf se třemi vrcholy. Úloha (12.27) 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 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 (12.25). Zaokrouhlené řešení je $\bar{\mathbf{x}} = (1, 1, 1)$ s optimální hodnotou 3, což je ve shodě s Větou 12.2. ◆

12.5.3 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 (12.27):

$$\max \sum_{i \in V} x_i \tag{12.29a}$$

$$\text{za podm. } x_i + x_j \leq 1, \quad \{i, j\} \in E \tag{12.29b}$$

$$x_i \in \{0, 1\}, \quad i \in V \tag{12.29c}$$

V LP relaxaci opět nahradíme podmínky $x_i \in \{0, 1\}$ nerovnostmi $0 \leq x_i \leq 1$.

Na rozdíl od úlohy na vrcholové pokrytí tato LP relaxace není příliš užitečná. LP relaxace má vždy přípustné řešení $x_i = \frac{1}{2}$ ($i \in V$). Tomu odpovídá hodnota účelové funkce $\sum_i x_i = \frac{1}{2}|V|$. Tedy optimální hodnota LP relaxace úlohy nemůže být menší než $\frac{1}{2}|V|$. Bohužel, optimální hodnota y^* nerelaxované úlohy (12.29) může libovolně-krát menší než $\frac{1}{2}|V|$. Např. pro každý úplný graf je $y^* = 1$ (protože každé dva vrcholy jsou spojeny hranou).

Mohli bychom doufat, že např. zaokrouhlením optimálního řešení relaxované úlohy získáme analogii Věty 12.2. Tyto snahy jsou ale beznadějně: bylo dokázáno, že úlohu (12.29) nelze aproximovat s jakoukoliv konstantní (tj. nezávislou na velikosti grafu) zárukou. Přesněji, pro každé číslo $\varepsilon > 0$ je nalezení přípustného řešení $\bar{x}_i \in \{0, 1\}$ ($i \in V$) splňujícího $\sum_{i \in V} \bar{x}_i \geq y^*/\varepsilon$ stejně těžké, jako nalezení optimálního řešení (úloha je tzv. *APX-těžká*).

12.6 Cvičení

12.1. 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) $\mathbf{c} = (0, 0, -1)$.

12.2. Následující úlohy nejprve převed'te na rovnicový tvar (tj. tvar s nezápornými proměnnými a omezeními typu lineární rovnice, viz §12.1). 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 $\mathbf{u}, \mathbf{P}, \mathbf{q}, \mathbf{r}$).

$$\begin{aligned} \text{a) } \quad & \min \quad 2x_1 \quad - 3x_3 + x_4 \\ \text{za podmíněk} \quad & x_1 - x_2 - x_3 \quad \geq 0 \\ & -x_1 + 2x_2 - 3x_3 \quad \leq 5 \\ & 2x_1 - x_2 - x_3 + 2x_4 = 6 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

b) lineární program (12.11)

c) lineární program (12.16)

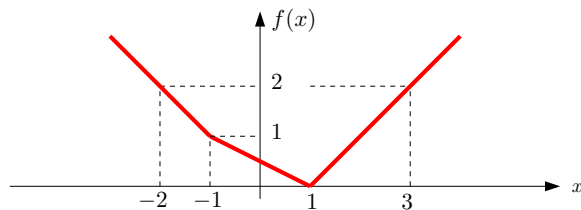
d) lineární program (12.18)

12.3. 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.

- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T\mathbf{x} \leq 1\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -1 \leq \mathbf{1}^T\mathbf{x} \leq 1\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T\mathbf{x} = k\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T\mathbf{x} \leq k\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, 0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1\}$
- (\star) $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, -\mathbf{y} \leq \mathbf{x} \leq \mathbf{y}, \mathbf{1}^T\mathbf{y} = k, \mathbf{y} \leq \mathbf{1}\}$
- $\min\{\mathbf{1}^T\mathbf{u} + \mathbf{1}^T\mathbf{v} \mid \mathbf{u}, \mathbf{v} \geq \mathbf{0}, \mathbf{u} - \mathbf{v} = \mathbf{c}\}$
- (\star) $\min\{\mathbf{a}^T\mathbf{u} + \mathbf{b}^T\mathbf{v} \mid \mathbf{u}, \mathbf{v} \geq \mathbf{0}, \mathbf{u} - \mathbf{v} = \mathbf{c}\}$, kde předpokládáme $\mathbf{a} \geq \mathbf{b}$

12.4. Pokuste se úlohy transformovat na LP. Pokud to nedokážete, vysvětlete proč.

- $\min\{|x_1| + |x_2| \mid x_1, x_2 \in \mathbb{R}, 2x_1 - x_2 \geq 1, -x_1 + 2x_2 \geq 1\}$
- $\max\{|x_1 - c_1| + \dots + |x_n - c_n| \mid x_1, \dots, x_n \in \mathbb{R}, a_1x_1 + \dots + a_nx_n \geq b\}$
- $\max\{\mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}, |\mathbf{d}^T\mathbf{x}| \leq 1, \mathbf{x} \geq \mathbf{0}\}$
- $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{l=1}^L \max_{k=1}^K (\mathbf{c}_{kl}^T\mathbf{x} + d_{kl})$
- $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^m f(\mathbf{a}_i^T\mathbf{x} - b_i)$, kde funkce f je definována obrázkem



- f) $\min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{Ax} - \mathbf{b}\|_\infty \leq 1 \}$
g) $\min\{ \|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} = \mathbf{b} \}$
h) $\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{Ax} - \mathbf{b}\|_1 + \|\mathbf{x}\|_\infty)$ (což se dá napsat také jako $\min\{ \|\mathbf{Ax} - \mathbf{b}\|_1 + \|\mathbf{x}\|_\infty \mid \mathbf{x} \in \mathbb{R}^n \}$)
i) $(\star) \min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{Ax} - \mathbf{b}\|_1 \leq 1 \}$

12.5. Máme algoritmus (černou skříňku) na řešení LP, kterou můžeme zavolat i vícekrát. S pomocí tohoto algoritmu vyřešte úlohu $\max\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}$.

12.6. Dokažte nebo vyvrat'te následující rovnosti. Zde $\mathbf{c} \in \mathbb{R}^n$ a $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou dány, $\|\cdot\|$ je libovolná norma, a optimalizuje se přes proměnné $\mathbf{x} \in \mathbb{R}^n$.

- 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) $\max\{ \|\mathbf{Ax}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1 \} = \max\{ \|\mathbf{Ax}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1 \}$

12.7. Dokažte, že

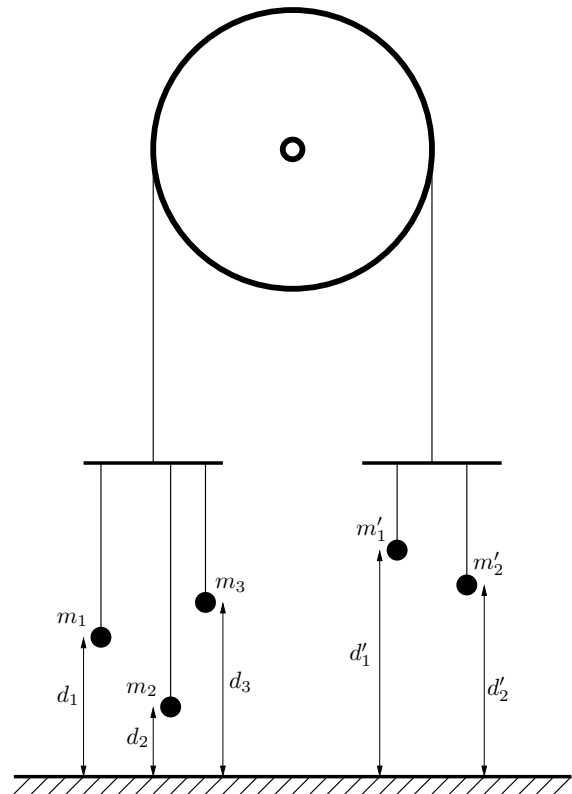
- a) $\min\{ \mathbf{1}^T (\mathbf{u} + \mathbf{v}) \mid \mathbf{u} - \mathbf{v} \in X, \mathbf{u}, \mathbf{v} \geq \mathbf{0} \} = \min_{\mathbf{x} \in X} \|\mathbf{x}\|_1,$
b) $\min\{ \mathbf{1}^T \mathbf{u} \mid \mathbf{u} - \mathbf{v} \in X, \mathbf{u}, \mathbf{v} \geq \mathbf{0} \} = \min_{\mathbf{x} \in X} \mathbf{x}_+$ kde značíme $\mathbf{x}_+ = \sum_{i=1}^n \max\{0, x_i\},$

kde $X \subseteq \mathbb{R}^n$ a předpokládáme, že všechny zmíněné úlohy mají optimální řešení.

12.8. Uvažujme Příklad 1.10 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.

12.9. 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.

12.10. 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.

12.11. 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.

12.12. Firma na výrobu kánoí má 120 zaměstnanců, z nichž každý pracuje maximálně 30 hodin týdně. Polovina zaměstnanců pracuje v truhlářské dílně, 20 zaměstnanců pracuje v dílně na zpracování plastů a zbytek v kompletační dílně. Firma vyrábí dva typy kánoí: standardní kánoe s čistým ziskem 7 EUR za kus a luxusní kánoe s čistým ziskem 10 EUR za kus. Na výrobu jedné standardní kánoe je třeba 4.5 hodiny práce v truhlářské dílně a dvě hodiny v každé ze zbylých dvou dílen. Jedna luxusní kánoe vyžaduje 5 hodin práce v

truhlárně, hodinu v dílně na plasty a 4 hodiny kompletace. Průzkum trhu odhalil, že ne méně než $1/3$ a ne více než $2/3$ vyrobených kánoí by měly být luxusní. Kolik kterých kánoí má firma týdně vyrobit, aby byl její čistý zisk maximální? Formalizujte jako optimalizační úlohu, kterou už ale neřešte.

12.13. Spočtete limitu $\lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \lim_{p \rightarrow \infty} (|x_1|^p + \dots + |x_n|^p)^{1/p}$.

12.14. Máme dvě konečné množiny $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ a $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ bodů z \mathbb{R}^d . Hledáme

- 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 jedné straně nadroviny a všechny body z druhé množiny jsou na druhé straně). Napište jako lineární program.

12.15. 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 $\mathbf{y} \in \mathbb{R}^{n-k}$)
- d) $f(\mathbf{x}) = \max_{i=1}^n x_i - \min_{i=1}^n x_i$

12.16. Dokažte, že p -norma splňuje axiomy vektorové normy.

12.17. Máme m nadrovin v \mathbb{R}^n , kde i -tá nadrovina je množina $H_i = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} = b_i\}$ kde $\mathbf{a}_i \neq \mathbf{0}$. Označme $d(H_i, \mathbf{x}) = \min_{\mathbf{y} \in H_i} \|\mathbf{x} - \mathbf{y}\|_2$ vzdálenost bodu $\mathbf{x} \in \mathbb{R}^n$ od i -té nadroviny. Hledáme bod \mathbf{x} , který minimalizuje číslo (a) $\sum_{i=1}^m d(H_i, \mathbf{x})$, (b) $\max_{i=1}^m d(H_i, \mathbf{x})$. Formulujte jako lineární program.

12.18. Hledáme \mathbf{x} , které minimalizuje výraz $f(\mathbf{A}\mathbf{x} - \mathbf{b})$, kde (a) $f(\mathbf{r}) = \|\mathbf{r}\|_\infty = \max_i |r_i|$, (b) $f(\mathbf{r}) = \|\mathbf{r}\|_1 = \sum_i |r_i|$, (c) $f(\mathbf{r}) = \sum_i \max\{\theta - r_i, 0, \theta + r_i\}$ kde $\theta > 0$ je daná konstanta. Formulujte jako lineární programy. Když bychom tyto úlohy interpretovali jako prokládání bodů regresní funkcí (jako v §12.4.3), jaký je geometrický význam těchto úloh?

12.19. (*) 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

12.2.b) Úloha už je v rovnicovém tvaru, takže v prvním kroku nemusíme dělat nic. Úloha jde napsat v maticovém tvaru (12.12), to ale není požadovaný tvar, protože proměnné jsou soustředěné do matice \mathbf{X} a nikoliv do vektoru. Převod do požadovaného tvaru je jasný z Příkladu .c.

12.3.a) Postupujeme podobně jako v Příkladu 12.4: nejdřív uhodneme optimální řešení a pak dokážeme jeho optimalitu tak, že ukážeme, že každé jiné řešení jde změnit (bez porušení přípustnosti) tak, že kritérium se zlepšit. Zde tedy uhodneme, že v optimu je $x_i = 0$ kdykoliv $c_i < 0$ a $x_i = 1$ kdykoliv $c_i > 0$. Kdyby totiž např. bylo $x_1 > 0$ a $c_1 < 0$, tak bychom mohli kritérium zlepšit položením $x_1 = 0$. Optimální hodnota odpovídající uhodnutému optimálnímu řešení je $\sum_i \max\{c_i, 0\}$.

12.3.b) Postupujeme obdobně. Optimální hodnota je $\sum_{i=1}^n |c_i|$.

12.3.c) $\max_{i=1}^n \max\{0, c_i\} = \max\{0, \max_{i=1}^n c_i\}$

12.3.d) Když $c_1 = c_2 = \dots = c_n = a$, tak optimální hodnota je $|a|$. Jinak je úloha neomezená.

12.3.g) Převeďte na jinou úlohu substitucí $y_i = x_i - x_{i-1}$ a pak postupujte jako v předešlých úlohách.

- 12.3.i) Dokažte nejprve, že v optimu je buď $\mathbf{u} = \mathbf{0}$ nebo $\mathbf{v} = \mathbf{0}$.
- 12.4.a) $\min\{z_1 + z_2 \mid x_1, x_2, z_1, z_2 \in \mathbb{R}, 2x_1 - x_2 \geq 1, -x_1 + 2x_2 \geq 1, x_1 \leq z_1, x_2 \leq z_2, -x_1 \leq z_1, -x_2 \leq z_2\}$
- 12.4.b) Nejde.
- 12.4.c) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{d}^T \mathbf{x} \leq 1, -\mathbf{d}^T \mathbf{x} \leq 1, \mathbf{x} \geq \mathbf{0}\}$
- 12.4.d) $\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{c}_{kl}^T \mathbf{x} + d_{kl} \leq z_l (\forall k, l), \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^L\}$ (analogické §12.1.1)
- 12.4.f) $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -\mathbf{1} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{1}\}$
- 12.4.h) $\min\{\mathbf{1}^T \mathbf{y} + z \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, z \in \mathbb{R}, -\mathbf{y} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{y}, -z\mathbf{1} \leq \mathbf{x} \leq z\mathbf{1}\}$
- 12.4.i) $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^m, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}, \mathbf{1}^T \mathbf{z} \leq 1\}$. Správnost tohoto převodu bychom museli dokázat podobně jako v Příkladu 12.4 – proved'te! Všimněte si také, že $-\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}$ implikuje $\mathbf{z} \geq \mathbf{0}$, tedy bychom mohli přidat podmínku $\mathbf{z} \geq \mathbf{0}$ a úloha by se nezměnila.
- 12.5. Postupem uvedeným v §12.1 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 $B = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.
- 12.6. Inspirujte se úvahou v §12.1.1.
- 12.7.a) Dokažte nejprve, že v optimu je buď $\mathbf{u} = \mathbf{0}$ nebo $\mathbf{v} = \mathbf{0}$. Pak např. uvažujte úvahu v §12.1.1 a použijte vhodnou substituci.
- 12.8. Stejná úloha jako v Příkladu 1.10, 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 §12.1.1.
- 12.9. 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 §12.1.1.
- 12.10.a) $E(x) = \sum_{i=1}^n m_i \max(d_i + x, 0) + \sum_{i=1}^{n'} m'_i \max(d'_i - x, 0)$
- 12.10.b) $\min\{\sum_{i=1}^n m_i z_i + \sum_{i=1}^{n'} m'_i z'_i \mid x, z_i, z'_i \in \mathbb{R}, z_i \geq d_i + x, z'_i \geq d'_i - x, z_i \geq 0, z'_i \geq 0\}$
- 12.12. $\max 7s + 10l$ z.p. $4.5s + 5l \leq 60 \cdot 30, 2s + l \leq 20 \cdot 30, 2s + 4l \leq 40 \cdot 30, (s + l)/3 \leq l \leq 2(s + l)/3, s, l \geq 0$ a navíc $s, l \in \mathbb{Z}$.
- 12.13. 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.
- 12.14.a) Podmínku oddělení napište jako soustavu lineárních nerovnic, což je vlastně LP s konstantní (nulovou) účelovou funkcí. Viz Příklad 17.12.
- 12.15.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 $\mathbf{x} = (-1, 0, 0, \dots)$.
- 12.19. Dle Věty 12.1 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 13

Konvexní množiny a mnohostěny

Zde si řekneme více o geometrii lineárního programování. Ke konvexním množinám, zásadnímu to pojmu v optimalizaci, se navíc vrátíme v pozdějších kapitolách.

13.1 Konvexní množiny

Množina $X \subseteq \mathbb{R}^n$ se nazývá **konvexní**, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies (1 - \alpha)\mathbf{x} + \alpha\mathbf{y} \in X. \quad (13.1)$$

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 :



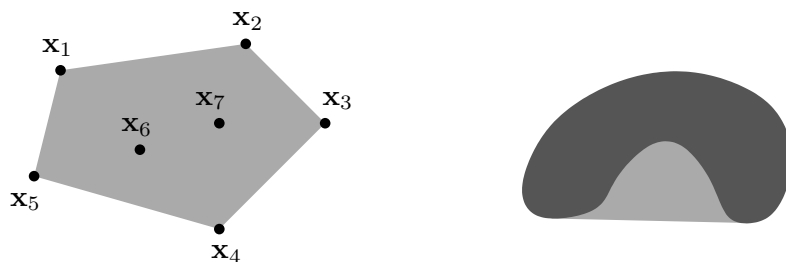
Konvexní kombinace bodů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je jejich lineární kombinace $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k$ taková, že $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Lze dokázat, že množina je konvexní právě tehdy, když je uzavřená na konvexní kombinace (neboli každá konvexní kombinace bodů 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 bodů \mathbf{x}, \mathbf{y} , neboť $(1 - \alpha) + \alpha = 1$, $1 - \alpha \geq 0$, $\alpha \geq 0$.

Konvexní obal bodů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich konvexních kombinací, značíme

$$\text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} = \{\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k \mid \alpha_1 + \dots + \alpha_k = 1, \alpha_1, \dots, \alpha_k \geq 0\}. \quad (13.2)$$

Jak ale definovat konvexní obal množiny s *nekonečným* počtem prvků (např. na pravém obrázku výše)? Nelze použít definice (13.2), neboť není jasné, co znamená součet $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k$ pro nekonečný počet bodů. **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 13.1. 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 $X \cap Y$. ■

Sjednocení konvexních množin ale nemusí být konvexní množina.

13.2 Čtyři kombinace, čtyři obaly

Konvexní kombinace je lineární kombinace, jejíž koeficienty splňují omezení $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Všimněte si, že když vynecháme druhé omezení, dostaneme afinní kombinaci (viz §3.3). Podle toho, které ze dvou omezení vyžadujeme, dostaneme čtyři druhy kombinací. Udělejme si v nich nyní pořádek.

Vážený součet $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k$ vektorů¹ $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ 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 na

- lineární kombinace, se nazývá **lineární podprostor**.
- afinní kombinace, se nazývá **afinní podprostor**.
- nezáporné kombinace, se nazývá **konvexní kužel**.
- konvexné kombinace, 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 13.1. 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. ◆

¹Ve světle poznámky v §3.3 tyto vektory také nazýváme *body*, vystupují-li v afinní nebo konvexní kombinaci.

13.3 Konvexní mnohostěny

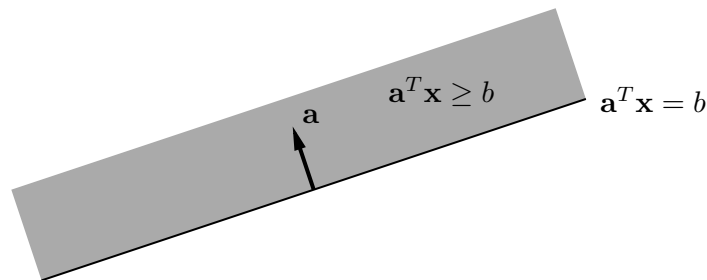
Zopakujme (viz §3.3), že *nadrovina* v \mathbb{R}^n je afinní podprostor dimenze $n-1$. Je to tedy množina

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}, \quad (13.3)$$

kde $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Požadujeme $\mathbf{a} \neq \mathbf{0}$, jinak by afinní podprostor (13.3) neměl dimenzi $n-1$ (a nebyl by tedy nadrovinou). Vektor \mathbf{a} je normála nadroviny a $|b|/\|\mathbf{a}\|$ je vzdálenost nadroviny od počátku (viz §4.6.2). **Poloprostor** (přesněji *uzavřený poloprostor*²) je množina

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \geq b\}. \quad (13.4)$$

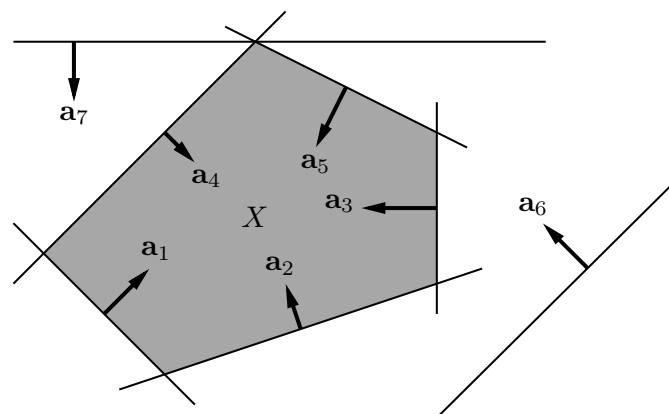
Hranice (viz §9.1) poloprostoru je nadrovina (13.3). Obrázek ilustruje tyto pojmy pro $n=2$:



Konvexní mnohostěn (krátce jen **mnohostěn**, angl. *polyhedron*) je průnik konečně mnoha uzavřených poloprostorů. Je to tedy množina

$$X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} \geq b_i \quad \forall i = 1, \dots, m\} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}, \quad (13.5)$$

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 13.1 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 §13.2, dimenze afinního podprostoru viz §3.3). Jinými slovy, je to největší počet afinně nezávislých bodů, které lze do mnohostěnu umístit, minus jedna.

²Otevřený poloprostor je množina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} > b\}$.

Příklad 13.2. Množina X z Příkladu 12.1 je konvexní mnohostěn. ◆

Příklad 13.3. Příklady ‘jednoduchých’ konvexních mnohostěů 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 $\{\mathbf{x} \in \mathbb{R}^n \mid b_1 \leq \mathbf{a}^T \mathbf{x} \leq b_2\}$ (průnik dvou poloprostorů s opačnými normálami)
- 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 n -rozměrný mnohostěn který je konvexním obalem svých $n + 1$ vrcholů
- standardní simplex $\{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = 1\}$ (tedy $(n - 1)$ -rozměrný konvexní mnohostěn, jehož vrcholy jsou vektory $\mathbf{e}_1, \dots, \mathbf{e}_n$ standardní báze \mathbb{R}^n)
- křížový polytop $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \leq 1\}$ (pro $n = 3$ pravidelný osmistěn) ◆

Příklad 13.4. Příklady konvexních množin, které nejsou mnohostěy:

- 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 $(0, 1)$ ◆

13.3.1 Extremální body

Bod $\mathbf{x} \in X$ se nazývá **extremální bod** konvexní množiny X , jestliže neexistují dva různé body z X takové, že \mathbf{x} je střed úsečky spojující tyto dva body, tj. jestliže platí implikace

$$\mathbf{x}_1, \mathbf{x}_2 \in X, \quad \mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) \implies \mathbf{x}_1 = \mathbf{x}_2. \quad (13.6)$$

Pro mnohostěn (13.5) a pro množinu $I \subseteq \{1, \dots, m\}$ budeme symbolem \mathbf{A}_I označovat³ 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 $\mathbf{a}_i^T \mathbf{x} = b_i \forall i \in I$ (pro $I = \emptyset$ definujeme soustavu jako prázdnou a její řešení je libovolné $\mathbf{x} \in \mathbb{R}^n$).

³Zde je drobná formální nesrovnalost: I je množina, a tedy na pořadí jejích prvků nezáleží, ale na pořadí řádků matice \mathbf{A}_I záleží. Abychom byli přesní, musíme proto předpokládat, že I je *uspořádaná množina*, kde její uspořádání je zděděné od množiny $\{1, \dots, m\}$, a řádky matice \mathbf{A}_I jsou seřazené tímto uspořádáním.

Věta 13.2. Pro každý bod \mathbf{x} mnohostěnu (13.5) jsou následující výroky ekvivalentní:

- Bod \mathbf{x} je extrémální bod mnohostěnu.
- Existuje neprázdná množina $I \subseteq \{1, \dots, m\}$ tak, že $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ a sloupce matice \mathbf{A}_I jsou lineárně nezávislé.

Důkaz. Necht' X označuje mnohostěn (13.5) a $\mathbf{x} \in X$ je jeho extrémální bod. Označme jako $I = \{i \mid \mathbf{a}_i^T \mathbf{x} = b_i\}$ množinu indexů nerovností v definici (13.5), 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 3.11), 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 (13.6) by proto \mathbf{x} nebyl extrémální bod.

Necht' I je taková, že $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ a sloupce \mathbf{A}_I jsou lineárně nezávislé. Necht' $\mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$ pro nějaké $\mathbf{x}_1, \mathbf{x}_2 \in X$, tedy

$$\begin{aligned}\mathbf{A}_I \mathbf{x}_1 &\geq \mathbf{b}_I, \\ \mathbf{A}_I \mathbf{x}_2 &\geq \mathbf{b}_I, \\ \mathbf{A}_I \mathbf{x} &= \frac{1}{2}(\mathbf{A}_I \mathbf{x}_1 + \mathbf{A}_I \mathbf{x}_2) = \mathbf{b}_I.\end{aligned}$$

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

$$\alpha_1 \geq \frac{1}{2}(\alpha_1 + \alpha_2) \leq \alpha_2 \implies \alpha_1 = \alpha_2. \quad (13.7)$$

Tedy platí $\mathbf{A}_I \mathbf{x}_1 = \mathbf{A}_I \mathbf{x}_2$. Protože však \mathbf{A}_I má lineárně nezávislé sloupce, plyne z toho $\mathbf{x}_1 = \mathbf{x}_2$ (dle Věty 3.1). Tedy \mathbf{x} je extrémální bod. ■

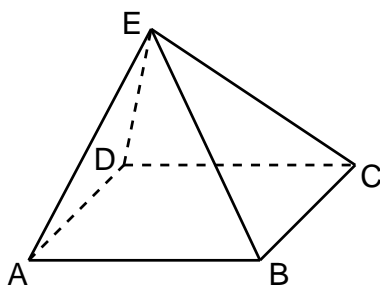
Připomeňme, že lineární soustava má právě jedno řešení právě tehdy, když má aspoň jedno řešení a její matice má lineárně nezávislé sloupce. Věta tedy vlastně říká, že $\mathbf{x} \in \mathbb{R}^n$ je extrémální bod mnohostěnu, právě když soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ má právě jedno 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 vypsát všechny extrémální 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 extrémální bod.

Příklad 13.5. Mějme mnohostěn $\{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, x + y \geq 1\}$ (nakreslete si ho!). Tedy v (13.5) máme $n = 2, m = 3$ a

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Zkoumejme podmnožiny $I \subseteq \{1, 2, 3\}$. Pro všechny jednoprvkové podmnožiny ($|I| = 1$) jistě má matice \mathbf{A}_I lineárně závislé sloupce. Pro všechny dvouprvkové podmnožiny ($\{1, 2\}, \{1, 3\}$ a $\{2, 3\}$) má matice \mathbf{A}_I lineárně nezávislé sloupce. Např. pro $I = \{1, 2\}$ je řešením soustavy $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ bod $\mathbf{x} = (x, y) = (0, 0)$, ten ale nespĺňuje $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ (tj. není v mnohostěnu), proto to není extrémální bod. Pro $I = \{1, 3\}$ má soustava jediné řešení $(x, y) = (0, 1)$, které patří do mnohostěnu a tudíž je to jeho extrémální bod. Pro $I = \{1, 2, 3\}$ nemá soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ žádné řešení. ♦

Příklad 13.6. Necht' mnohostěn (13.5) je pyramida v \mathbb{R}^3 na obrázku:

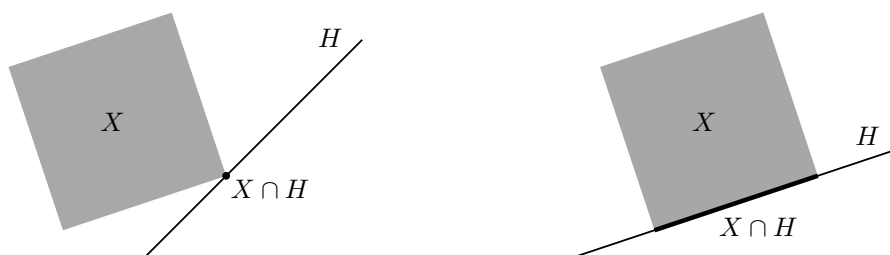


Tento mnohostěn je průnikem pěti poloprostorů, tedy $n = 3$ a $m = 5$. Necht' omezení $\mathbf{a}_i^T \mathbf{x} \geq b_i$ pro $i = 1, 2, 3, 4, 5$ je poloprostor, jehož hranicí je nadrovina určená po řadě body ABCD, ABE, BCE, CDE, ADE. Pro $I = \{1\}$ má soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ řešení ale sloupce matice \mathbf{A}_I jsou lineárně závislé, tedy soustava má nekonečně mnoho řešení, a proto žádné její řešení není extrémální bod. Pro $I = \{1, 2, 3, 4, 5\}$ nemá soustava žádné řešení. Pro $I = \{1, 2, 3\}$ má soustava řešení a navíc sloupce matice \mathbf{A}_I jsou lineárně nezávislé, tedy soustava má právě jedno řešení. Toto řešení navíc splňuje soustavu $\mathbf{A} \mathbf{x} \geq \mathbf{b}$, tedy je to extrémální bod (vrchol B). ♦

Protože podmnožin I je konečně mnoho, má každý konvexní mnohostěn konečně mnoho extrémálních bodů. Počet extrémálních bodů mnohostěnu popsaného soustavou lineárních nerovnic ovšem může být exponenciální funkcí počtu nerovnic (tedy může být astronomicky velký pro nepříliš velký počet nerovnic). Tomu se není co divit, protože počet všech podmnožin množiny $\{1, \dots, m\}$ je 2^m . Např. hyperkrychle $\{\mathbf{x} \in \mathbb{R}^n \mid -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$ je popsána $2n$ lineárními nerovnicemi ale má 2^n extrémálních bodů. Není tomu tak vždy, např. simplex (viz Příklad 13.3) je popsán $n + 1$ lineárními nerovnicemi a má $n + 1$ extrémálních bodů.

13.3.2 Stěny mnohostěnu

Opěrná nadrovina konvexní množiny $X \subseteq \mathbb{R}^n$ je nadrovina $H = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$ 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 13.1 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 extrémální bod. V následujícím textu budeme používat pojem extrémální bod (se kterým se pracuje lépe než s ekvivalentním pojmem vrchol).

Každým hraničním bodem (viz §9.1) 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.

13.3.3 Extrémy lineární funkce na mnohostěnu

Lema 13.1. *Je-li H opěrná nadrovina konvexní množiny X , pak každý extrémální bod množiny $X \cap H$ je extrémální bod množiny X .*

Důkaz. Necht' $H = \{ \mathbf{x} \mid \mathbf{a}^T \mathbf{x} = b \}$ a $\mathbf{a}^T \mathbf{x} \geq b$ pro všechna $\mathbf{x} \in X$. Necht' $\mathbf{x} \in X \cap H$ není extrémální 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 (13.7) 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 extrémální bod mnohostěnu $X \cap H$. ■

Věta tedy říká, že extrémální bod stěny mnohostěnu je zároveň extrémální bod mnohostěnu.

Ne každý mnohostěn má extrémální 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 13.3. *Pro každý neprázdný konvexní mnohostěn jsou následující výroky ekvivalentní:*

- Mnohostěn má aspoň jeden extrémální bod.
- Mnohostěn neobsahuje přímku.

Důkaz. Nejdříve dokážeme, že když mnohostěn X obsahuje přímku, tak nemá ani jeden extrémální 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 (13.5), 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 extrémální bod.

Nyní dokážeme, že jestliže mnohostěn neobsahuje přímku, pak má aspoň jeden extrémální 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 extrémální bod. Ten je dle Lematu 13.1 extrémální bod mnohostěnu X . ■

Věta 13.4. 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 extrémálních 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 13.3 má aspoň jeden extrémální bod. Dle Lematu 13.1 je tento bod extrémální bod mnohostěnu X . ■

Hledejme nyní minimum lineární funkce na mnohostěnu, neboli řešme lineární program

$$\min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b} \}.$$

Jestliže mnohostěn neobsahuje přímku a funkce na něm má minimum, pak dle Věty 13.4 stačí jednoduše projít všechny extrémální 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 extrémální body, neboť jich je příliš mnoho. V příští kapitole popíšeme *simplexový algoritmus*, který téměř všechny tyto potíže řeší.

13.4 Cvičení

13.1. Odpovězte, zda následující množiny jsou konvexní a odpověď dokažte z definice konvexní množiny:

- a) interval $[a, b]$
- 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) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{C}\mathbf{x} = \mathbf{d} \}$
- f) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{x} \leq 1 \}$
- g) \mathbb{Z} (množina celých čísel)
- h) $\{ \mathbf{x} \in \mathbb{R}^n \mid \max\{x_1, \dots, x_n\} \geq 0 \}$
- i) $\{ \mathbf{C}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b} \}$ (lineární zobrazení konvexního mnohostěnu)

13.2. 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) $\{(1, 1), (1, 2), (3, 1)\}$

13.3. (★) Zobrazení $f: 2^V \rightarrow 2^V$ (kde V je libovolná množina a 2^V značí množinu všech podmnožin množiny V) se nazývá *uzávěr* (angl. *closure*) na množině V , jestliže pro každé $X, Y \subseteq V$ platí

- $X \subseteq f(X)$ (extensivita),
- $f(f(X)) = f(X)$ (idempotence),
- $X \subseteq Y \Rightarrow f(X) \subseteq f(Y)$ (monotonicita).

Dokažte, že operace lineárního, afinního, nezáporného a konvexního obalu jsou uzávěry. Tedy dokažte tři vlastnosti výše pro $V = \mathbb{R}^n$ a např. $f(X) = \text{conv } X$.

13.4. (★) Navrhněte algoritmus na nalezení konvexního obalu konečné množiny bodů v rovině. Výstupem bude uspořádný seznam bodů, které tvoří extrémální body konvexního obalu.

13.5. Nakreslete lineární, afinní, nezáporný a konvexní obal náhodně zvolených k vektorů v \mathbb{R}^n pro všech devět případů $k, n \in \{1, 2, 3\}$.

13.6. Bude Věta 13.1 platit, pokud v ní výraz ‘konvexní množina’ nahradíme výrazem ‘lineární podprostor’ (příp. ‘afinní podprostor’, ‘konvexní kužel’)? Odpověď dokažte.

13.7. Jsou následující množiny konvexní mnohostěny? Zápornou odpověď odůvodněte. Kladnou odpověď dokažte tak, že množinu napíšete jako množinu řešení soustavy konečně mnoha lineárních nerovnic (tj. jako průnik konečně mnoha 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

13.8. Mějme body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. 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. Sjednocení hranic těchto množin se nazývá *Voronoiův diagram*. Nakreslete si ho pro $n = 2$ a pro tři případy $m \in \{2, 3, 4\}$ pro různé konfigurace bodů $\mathbf{a}_1, \dots, \mathbf{a}_m$.

13.9. Hledáme největší (hyper)kouli $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{c}\|_2 \leq r\}$, která se vejde do mnohostěnu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b}\}$. Zformulujte jako lineární program.

13.10. Chceme najít všechny extrémální body mnohostěnu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b}\}$.

- a) Udělejte pro mnohostěn

$$\mathbf{A} = \begin{bmatrix} -1 & -2 \\ 3 & -1 \\ -1 & 1 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -14 \\ 0 \\ -2 \\ 4 \end{bmatrix}.$$

- b) Napište program v Matlabu, který vypíše všechny extrémální body pro libovolné \mathbf{A}, \mathbf{b} .
- 13.11. Které z následujících mnohostěnů mají aspoň jeden extrémální bod? Kladnou odpověď dokažte nalezením libovolného jednoho extrémálního bodu. Zápornou odpověď dokažte nalezením libovolné přímky, která leží v mnohostěnu.
- a) všechny mnohostěny z Příkladu 13.3
 b) $\{ (x, y, z) \in \mathbb{R}^3 \mid x + y \leq 1, x \geq 0, y \geq 0 \}$
 c) $\{ (x, y) \in \mathbb{R}^2 \mid -1 \leq x + y \leq 1 \}$
- 13.12. (★) Pro mnohostěn $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b} \}$ dokažte:
- a) Má-li \mathbf{A} lineárně závislé sloupce (tj. $\text{rank } \mathbf{A} < n$), mnohostěn nemá extrémální bod.
 b) Má-li mnohostěn aspoň jeden extrémální bod, je $\text{rank } \mathbf{A} = n$.

Nápověda a řešení

- 13.1.a) Konvexní, protože pro libovolné $\alpha \in [0, 1]$ je $(1 - \alpha)a + \alpha b \in [a, b]$.
- 13.1.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.
- 13.1.c) Je konvexní.
- 13.1.e) Konvexní.
- 13.1.f) Konvexní.
- 13.1.g) Nekonvexní. Např. pro $x = 1, y = 2, \alpha = \frac{1}{2}$ číslo $(1 - \alpha)x + \alpha y = 1.5$ není celé.
- 13.1.h) Nekonvexní.
- 13.2.a) průnik poloprostorů a nadroviny, konvexní mnohostěn
- 13.2.b) Sféra, nekonvexní. Konv. obal je koule $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq 1 \}$.
- 13.2.c) koule bez hranice, konvexní
- 13.2.d) Graf paraboly, nekonvexní. Konvexní obal je $\{ (x, y) \in \mathbb{R}^2 \mid y \geq x^2 \}$.
- 13.2.e) Jedna větev grafu hyperboly, není konvexní. Konv. obal je $\{ (x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, xy \geq 1 \}$.
- 13.2.f) průnik dvou koulí, konvexní
- 13.2.g) Čtvrt kružnice, nekonvexní. Konv. obal je $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, x + y \geq 1 \}$.
- 13.2.h) Tři body v \mathbb{R} , nekonvexní protože např. $(-1 + 0)/2$ nepatří do množiny. Konv. obal je interval $[-1, 1]$.
- 13.2.i) Dva body v \mathbb{R}^2 , nekonvexní. Konv. obal je úsečka spojující ty dva body. Tuto úsečka máme napsat jako množinu řešení rovnic a nerovnic, což jde např. takto: $\{ (x, y) \in \mathbb{R}^2 \mid 1 \leq x \leq 2, x = y \}$.
- 13.2.j) Tři body v \mathbb{R}^2 , nekonvexní. Konv. obal je trojúhelník $\{ (x, y) \in \mathbb{R}^2 \mid x \geq 1, y \geq 1, x + 2y \leq 5 \}$.
- 13.7.a) Je to elipsoid, tedy je to konvexní množina. Mnohostěn je to jen pro $n = 1$ (v tom případě je to úsečka).
- 13.7.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 $\mathbf{Ax} = \mathbf{0}$, což už je průnik poloprostorů.
- 13.9. Maximalizujeme r za podmínek $\mathbf{Ac} - \mathbf{b} \geq r\mathbf{1}$ (tedy neznámé jsou $\mathbf{c} \in \mathbb{R}^n$ a $r \geq 0$), přičemž předpokládáme, že v každé nerovnici $\mathbf{a}_i^T \mathbf{x} \geq b_i$ soustavy $\mathbf{Ax} \geq \mathbf{b}$ máme $\|\mathbf{a}_i\| = 1$ (tedy každou nerovnici musíme nejdříve vydělit číslem $\|\mathbf{a}_i\|$).

- 13.10.a) Extremální body a odpovídající množiny $I \subseteq \{1, \dots, 4\}$:
- | | | | | |
|--------------|------------|------------|-------------------------------|------------|
| I | $\{1, 2\}$ | $\{1, 3\}$ | $\{2, 4\}$ | $\{3, 4\}$ |
| \mathbf{x} | $(2, 6)$ | $(6, 4)$ | $(\frac{4}{5}, \frac{12}{5})$ | $(2, 0)$ |
- 13.11.b) nemá extrémální bod
- 13.11.c) nemá extrémální bod
- 13.12.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 extrémální bod. Srov. Věta [13.3](#).

Kapitola 14

Simplexová metoda

Nyní popíšeme slavný algoritmus na řešení lineárních programů, **simplexovou metodu**. V minulé kapitole jsme ukázali, že má-li mnohostěn aspoň jeden extrémální bod, lineární funkce na něm nabývá minima aspoň v jednom extrémálním bodě. Tedy můžeme projít všechny extrémální body a najít ten s nejlepší účelovou funkcí. Ale extrémálních bodů může být mnoho. Simplexová metoda tento naivní algoritmus výrazně zrychluje tak, že postupuje od jednoho extrémálního bodu k dalšímu po hranách mnohostěnu tak, aby se účelová funkce zlepšovala nebo aspoň nezhoršovala. Ve zbytku kapitoly tento geometricky popsany nápad převedeme do algebry.

Pro začátek zapomeňme na účelovou funkci a zkoumejme pouze mnohostěn přípustných řešení LP. Místo mnohostěnu v obecném tvaru (13.5) algoritmus pracuje s mnohostěnem v rovnicovém tvaru (12.3)

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}, \quad (14.1)$$

kde navíc matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ má lineárně nezávislé řádky (tj. $\text{rank } \mathbf{A} = m$). Množina řešení soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ je tedy afinní podprostor dimenze nejvýše $n - m$.

Zkoumejme soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$ s dodatečnou podmínkou, že některé proměnné nastavíme na nulu. Tedy zvolíme množinu indexů $J \subseteq \{1, 2, \dots, n\}$ a řešíme soustavu

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (14.2a)$$

$$x_j = 0 \quad \forall j \in \{1, \dots, n\} \setminus J. \quad (14.2b)$$

Kdy má tato soustava právě jedno řešení? Dle Věty 3.1 právě tehdy, když sloupce matice \mathbf{A} s indexy J jsou lineárně nezávislé a \mathbf{b} patří do jejich lineárního obalu. Aby sloupce J byly lineárně nezávislé, nutně musí být $|J| \leq m$. Když ovšem $|J| < m$, pak existuje množina $J' \supseteq J$ tak, že $|J'| = m$ a sloupce J' jsou lineárně nezávislé (tedy doplníme sloupce J na bázi prostoru $\text{rng } \mathbf{A}$, což lze díky Větě 3.2). Pak ale, dle Věty 3.1, bude řešení soustavy splňovat $x_j = 0$ pro všechna $j \in J' \setminus J$. Tedy soustava bude mít stejné řešení pro J i J' . Proto, chceme-li najít všechny body \mathbf{x} která jsou jediným řešením soustavy (14.2), stačí soustavu vyřešit jen pro množiny J velikosti $|J| = m$, pro které jsou sloupce J matice \mathbf{A} lineárně nezávislé.

Příklad 14.1. Necht' je soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ dána tabulkou (blokovou maticí)

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccccc|c} -1 & 1 & 3 & 1 & 0 & 2 & 1 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \\ -1 & 0 & 4 & 1 & 1 & 4 & 2 \end{array} \right]. \quad (14.3)$$

Sloupce $J = \{1, 4, 5\}$ matice \mathbf{A} jsou lineárně nezávislé (ověřte!). Soustava (14.2) vypadá takto

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \quad x_2 = x_3 = x_6 = 0 \quad (14.4)$$

a má jediné řešení $\mathbf{x} = (3, 0, 0, 4, 1, 0)$. Sloupce J tvoří regulární matici velikosti $m \times m$.

Sloupce $J = \{3, 4\}$ jsou lineárně nezávislé a soustava

$$\begin{bmatrix} 3 & 1 \\ 4 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \quad x_1 = x_2 = x_5 = x_6 = 0$$

má jediné řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$. Sloupce $J = \{3, 4, 5\}$ jsou lineárně nezávislé a soustava

$$\begin{bmatrix} 3 & 1 & 0 \\ 4 & 0 & 1 \\ 4 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \quad x_1 = x_2 = x_6 = 0$$

má jediné řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$, stejné jako pro $J = \{3, 4\}$. Sloupce $J = \{3, 4, 6\}$ jsou také lineárně nezávislé a soustava má tedy stejné řešení. \blacklozenge

Uvedená úvaha motivuje zavedení následujících pojmů. Množina $J \subseteq \{1, \dots, n\}$ se nazývá **báze**, pokud $|J| = m$ a sloupce J matice \mathbf{A} jsou lineárně nezávislé. Proměnným x_j pro $j \in J$ říkáme **bázové proměnné**. Bod \mathbf{x} je **bázové řešení** příslušné bázi J , jestliže splňuje (14.2). Bázové řešení \mathbf{x} je **přípustné** (tedy leží v mnohostěnu), pokud $\mathbf{x} \geq \mathbf{0}$. Bázové řešení \mathbf{x} je **degenerované**, pokud má více než $m - n$ nulových složek. Protože matice \mathbf{A} má hodnost m , existuje aspoň jedna báze. Každé bázi přísluší právě jedno bázové řešení. Je-li ovšem bázové řešení degenerované, přísluší více než jedné bázi (v Příkladu 14.1 bázové řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$ přísluší bázím $\{3, 4, 5\}$ a $\{3, 4, 6\}$).

Mnohostěn (14.1) neobsahuje přímku, protože už jeho nadmnožina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}\}$ neobsahuje přímku (proč?). Tedy jestliže je mnohostěn (14.1) neprázdný, má dle Věty 13.3 aspoň jeden extrémální bod. Věta 13.2, vyslovená pro mnohostěn v obecném tvaru (13.5), jde vyslovit i pro mnohostěn v rovnicovém tvaru (14.1):

Důsledek 14.1. *Bod mnohostěnu (14.1) je extrémální, právě když je to přípustné bázové řešení.*

Důkaz. Mnohostěn (14.1) lze napsat ve tvaru (13.5) jako $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}'\mathbf{x} \geq \mathbf{b}'\}$, kde

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \\ \mathbf{I}_n \end{bmatrix} \in \mathbb{R}^{(2m+n) \times n}, \quad \mathbf{b}' = \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{2m+n}.$$

Máme dokázat, že pro každé $\mathbf{x} \in \mathbb{R}^n$ jsou následující dva výroky ekvivalentní:

1. $\mathbf{A}'\mathbf{x} \geq \mathbf{b}'$ a existuje množina $I \subseteq \{1, \dots, 2m+n\}$ tak, že \mathbf{x} je jediným řešením soustavy $\mathbf{A}'_I \mathbf{x} = \mathbf{b}'_I$.
2. $\mathbf{x} \geq \mathbf{0}$ a existuje množina $J \subseteq \{1, \dots, n\}$ tak, že \mathbf{x} je jediným řešením soustavy (14.2).

Implikaci $1 \Leftarrow 2$ dokážeme položením $I = \{1, \dots, 2m\} \cup \{2m+j \mid j \in \{1, \dots, n\} \setminus J\}$. Implikaci $1 \Rightarrow 2$ dokážeme položením $J = \{j \in \{1, \dots, n\} \mid 2m+j \notin I\}$. Rozmyslete, že to tak je! ■

Dvě báze nazveme **sousední**, pokud mají $m - 1$ společných prvků (např. báze $\{3, 4, 5\}$ a $\{3, 4, 6\}$ v Příkladu 14.1). Lze ukázat (důkaz vynecháme), že dvojice sousedních bází odpovídají buď jedinému (degenerovanému) extrémálnímu bodu nebo dvojici extrémálních bodů spojených hranou. Simplexová metoda přechází mezi sousedními bázemi tak, že bázová řešení jsou stále přípustná a účelová funkce se zlepšuje nebo aspoň nezhoršuje.

V §14.1 popíšeme stavební kameny metody, které pak v §14.2 spojíme do celého algoritmu.

14.1 Stavební kameny algoritmu

14.1.1 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 $\mathbf{Ax} = \mathbf{b}$: libovolný řádek tabulky $[\mathbf{A} \ \mathbf{b}]$ můžeme vynásobit nenulovým číslem a můžeme k němu přičíst lineární kombinaci ostatních řádků. Tyto úpravy nemění množinu řešení soustavy.

Ukážeme, jak přejít od aktuální standardní báze J k sousední standardní bázi, tedy nějaký sloupec $j' \in J$ bázi opustí a nějaký sloupec $j \notin J$ do báze vstoupí. Necht' i je řádek, ve kterém je $a_{ij'} = 1$. Prvek (i, j) matice se nazývá **pivot** (angl. znamená *čep*). Necht' $a_{ij} \neq 0$. Chceme nastavit pivot a_{ij} na jedničku, vynulovat prvky nad i pod pivotem, a nezměnit přitom sloupce $J \setminus \{j'\}$. Toho se dosáhne těmito ekvivalentními řádkovými úpravami:

1. Vyděl řádek i číslem a_{ij} .
2. Pro každé $i' \neq i$ odečti $a_{i'j}$ -násobek řádku i od řádku i' .

Říkáme, že jsme provedli *ekvivalentní úpravu kolem pivotu* s indexy (i, j) .

Příklad 14.2. Mějme soustavu

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & \boxed{1} & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array} \right] \quad (14.5)$$

se (standardní) bází $J = \{1, 4, 5\}$. Vidíme ihned odpovídající bázové řešení, $\mathbf{x} = (3, 0, 0, 4, 1, 0)$.

Chceme nahradit bázový sloupec $j' = 1$ nebázovým sloupcem $j = 2$, tedy přejít k sousední bázi $\{2, 4, 5\}$. Máme $i = 2$, tedy pivot je prvek a_{22} (v tabulce orámován). Ekvivalentními řádkovými úpravami musíme docílit, aby pivot byl roven jedné a prvky nad ním a pod ním byly nulové. Při tom smíme změnit sloupec 1, ale sloupce 4 a 5 se změnit nesmějí. Toho se docílí vydělením řádku 2 číslem a_{22} (což zde nemá žádný efekt, protože náhodou $a_{22} = 1$) a pak přičtením vhodných násobků řádku 2 k ostatním řádkům. Výsledek:

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} -2 & 0 & 0 & 1 & 0 & 0 & -2 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \end{array} \right].$$

Nyní sloupce $\{2, 4, 5\}$ jsou standardní báze. ◆

14.1.2 Kdy je nové bázové řešení přípustné?

Uvedeným způsobem můžeme od aktuální báze přejít k libovolné sousední bázi. Přitom nové bázové řešení může nebo nemusí být přípustné. Je-li aktuální bázové řešení přípustné, jak poznáme, zda i nové bázové řešení bude přípustné?

Protože nenulové složky bázového řešení \mathbf{x} jsou rovny složkám vektoru \mathbf{b} , bázové řešení je přípustné právě tehdy, když $\mathbf{b} \geq \mathbf{0}$. Necht' v aktuální tabulce je $\mathbf{b} \geq \mathbf{0}$. 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 §14.1.1):

- 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:

$$a_{ij} > 0, \quad (14.6a)$$

$$\text{pro každé } i' \neq i \text{ platí } a_{i'j} \leq 0 \text{ nebo } \frac{b_i}{a_{ij}} \leq \frac{b_{i'}}{a_{i'j}}, \quad (14.6b)$$

kde 'nebo' je užito v nevylučovacím smyslu. Podmínka (14.6a) je zřejmá. Podmínka (14.6b) 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 14.3. V tabulce (14.5):

- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 2)$ nepovede k přípustnému bázovému řešení, neboť $a_{ij} = -1 < 0$, což porušuje podmínku (14.6a).
- Ekvivalentní úprava okolo pivotu $(i, j) = (2, 2)$ nepovede k přípustnému bázovému řešení, neboť pro $i' = 1$ je $a_{i'j} > 0$ a $\frac{3}{1} > \frac{4}{2}$, tedy podmínka (14.6b) je porušena.
- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 6)$ povede k přípustnému bázovému řešení. Podmínky (14.6) jsou splněny, neboť $a_{ij} = 2 > 0$ a $\frac{1}{2} \leq \frac{4}{4}$, $\frac{1}{2} \leq \frac{3}{2}$. ♦

14.1.3 Co když je celý sloupec nekladný?

Jestliže jsou všechny prvky v nějakém nebázovém sloupci j nekladné, víme z podmínky (14.6a), že tento sloupec se nemůže stát bázovým. Platí ale navíc, že souřadnice x_j bodu \mathbf{x} se může libovolně zvětšovat a bod \mathbf{x} přesto zůstane v mnohostěnu X . Tedy existuje polopřímka s počátkem v \mathbf{x} ležící celá v mnohostěnu X . Tedy mnohostěn X je neomezený.

Příklad 14.4. 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 bázi $\{1, 4, 5\}$. Pod tabulkou je napsáno bázové řešení \mathbf{x} . Když se x_2 bude libovolně zvětšovat, změnu lze kompenzovat současným zvětšováním bázových proměnných x_1, x_4, x_5 tak, že vektor \mathbf{Ax} zůstane nezměněn a tedy roven \mathbf{b} . Konkrétně, 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 $\mathbf{x} \geq \mathbf{0}$. ♦

14.1.4 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

$$\min\{\mathbf{c}^T \mathbf{x} - d \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}. \quad (14.7)$$

Úlohu budeme reprezentovat **simplexovou tabulkou**

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix}. \quad (14.8)$$

Přičtíme k účelovému řádku $[\mathbf{c}^T \quad d]$ libovolnou lineární kombinaci $\mathbf{y}^T [\mathbf{A} \quad \mathbf{b}]$ ostatních řádků $[\mathbf{A} \quad \mathbf{b}]$, 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

$$[\mathbf{c}^T \quad d] + \mathbf{y}^T [\mathbf{A} \quad \mathbf{b}] = [\mathbf{c}^T + \mathbf{y}^T \mathbf{A} \quad d + \mathbf{y}^T \mathbf{b}].$$

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 $\mathbf{Ax} = \mathbf{b}$.

14.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?

Necht' J je standardní báze. Přičtíme 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 bázové řešení \mathbf{x} je v nebázových sloupcích nulové, znamená to $\mathbf{c}^T \mathbf{x} = 0$. Tedy hodnota účelové funkce $\mathbf{c}^T \mathbf{x} - d$ v bázovém řešení \mathbf{x} je rovna jednoduše $-d$.

Navíc je snadno vidět, co udělá přechod k nové bázi s účelovou funkcí. Necht' j' je sloupec opouštějící bázi a j je sloupec vstupující do báze. Při přechodu k nové bázi se číslo $x_{j'}$ stane nulovým a číslo x_j se zvětší z nuly na kladné (nebo se nezmění). Protože $c_{j'} = 0$, číslo $\mathbf{c}^T \mathbf{x} - d$ při $c_j \geq 0$ stoupne (nebo se nezmění) a při $c_j \leq 0$ klesne (nebo se nezmění).

Příklad 14.5. 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 bázových sloupcích vynulujeme tak, že k účelovému řádku přičtíme 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 bázové řešení \mathbf{x} . Nyní je $\mathbf{c}^T \mathbf{x} = 0$, a tedy hodnota účelové funkce v bázovém řešení je $\mathbf{c}^T \mathbf{x} - d = -d = -3$.

Dejme tomu, že chceme přidat do báze nebázový sloupec 2 a vyloučit z ní některý z bázových sloupců $\{1, 4, 5\}$. Po tomto přechodu se x_2 stane kladné nebo zůstane nulové a jedna ze složek x_1, x_4, x_5 se vynuluje. Protože $c_1 = c_4 = c_5 = 0$, změna x_1, x_4, x_5 se na účelové funkci neprojeví a ta se změní o $c_2 x_2$. Kritérium tedy stoupne nebo zůstane stejné, protože $c_2 = 1 > 0$. ♦

Jestliže v některém sloupci j je $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i , proměnnou x_j můžeme libovolně zvětšovat (viz §14.1.3) a účelovou funkci libovolně zmenšovat. Úloha je tedy neomezená.

14.2 Základní algoritmus

Spojením popsaných stavebních kamenů dostaneme iteraci simplexové metody na řešení úlohy (14.7). Iterace přejde k sousední standardní bázi takové, že bázové řešení zůstane přípustné a účelová funkce se nezmění. Vstupem i výstupem iterace je simplexová tabulka (14.8) s těmito vlastnostmi:

- podmnožina sloupců \mathbf{A} je standardní báze J ,
- bázové řešení odpovídající této bázi je přípustné, $\mathbf{b} \geq \mathbf{0}$,
- složky vektoru \mathbf{c} v bázových sloupcích jsou nulové, $c_j = 0$ pro $j \in J$.

Iterace se provede takto:

1. Vyber index j pivotu tak, aby¹ $c_j < 0$ (§14.1.5).
2. Vyber index i pivotu podle podmínek (14.6). Z těchto podmínek plyne (promyslete!)

$$i \in \operatorname{argmin}_{i' | a_{i'j} > 0} \frac{b_{i'}}{a_{i'j}}, \quad (14.9)$$

kde $\operatorname{argmin}_{i' | a_{i'j} > 0}$ označuje, že se minimalizuje přes všechna i' splňující $a_{i'j} > 0$.

3. Udělej ekvivalentní úpravu okolo pivotu (i, j) (§14.1.1).
4. Ekvivalentní úpravou účelového řádku vynuluj c_j v novém bázovém sloupci j (§14.1.5).

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. Pak účelovou funkci nelze zlepšit a jsme v optimu²
- Existuje sloupec j , ve kterém $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i . Pak je úloha neomezená.

Výběr indexů (i, j) pivotu v krocích 1 a 2 nemusí být jednoznačný: může existovat více sloupců j s vhodným znaménkem c_j a více řádků i' může splňovat podmínky (14.6) (tedy množina $\{b_{i'}/a_{i'j} \mid i' = 1, \dots, m, a_{i'j} > 0\}$ může mít více minimálních prvků). Algoritmus, který vybírá jediný pivot z těchto možností, se nazývá **pivotové pravidlo**.

¹Když je lineární program ve tvaru maximalizace, samozřejmě ho nemusíte převádět do minimalizačního tvaru (14.7), stačí jen vybrat sloupec s $c_j > 0$.

²Pozor: aktuální bázové řešení může být optimální přesto, že nějaký koeficient c_j je záporný. V další iteraci vložíme sloupec j do báze, ale kvůli degeneraci může zůstat $x_j = 0$, tedy účelová funkce se nezmění (viz §14.1.5).

Příklad 14.6. Vyřešte lineární program (14.7) simplexovou metodou, když výchozí simplexová tabulka (14.8) 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 bázové 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 (14.9) 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 $J = \{1, 4, 6\}$, bázové 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 (14.9) 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 $(x_1, x_2, x_3, x_4, x_5, x_6) = (1, 2, 0, 0, 3, 0)$. \blacklozenge

Příklad 14.7. Necht' simplexová tabulka (14.8) je

$$\begin{array}{cccccc|c} -2 & 6 & 1 & 0 & 0 & 0 \\ \hline -1 & -1 & -1 & 1 & 0 & 2 \\ \boxed{2} & -1 & -2 & 0 & 1 & 1 \end{array}$$

Tabulka po první iteraci je

$$\begin{array}{cccccc|c} 0 & 5 & -1 & 0 & 1 & 1 \\ \hline 0 & -1.5 & -2 & 1 & 0.5 & 2.5 \\ 1 & -0.5 & -1 & 0 & 0.5 & 0.5 \end{array}$$

Podle multého řádku by další pivot měl být ve třetím sloupci. Ale čísla a_{i3} jsou všechna záporná (viz §14.1.3). 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. ♦

14.2.1 Cyklení

Zřídka se algoritmus může dostat do stavu, kdy cyklicky prochází stále stejnou množinou bází, které odpovídají jedinému degenerovanému bázovému řešení a tedy účelová funkce se nemění. Algoritmus tedy nikdy neskončí. Tomuto chování říkáme **cyklení** algoritmu.

Příklad 14.8. Zde je počáteční simplexová tabulka a dvě iterace simplexového algoritmu:

$$\begin{array}{cccccc|c} -2.3 & -2.15 & 13.55 & 0.4 & 0 & 0 & 0 \\ \hline \boxed{0.4} & 0.2 & -1.4 & -0.2 & 1 & 0 & 0 \\ -7.8 & -1.4 & 7.8 & 0.4 & 0 & 1 & 0 \\ \\ 0 & -1 & 5.5 & -0.75 & 5.75 & 0 & 0 \\ \hline 1 & 0.5 & -3.5 & -0.5 & 2.5 & 0 & 0 \\ 0 & \boxed{2.5} & -19.5 & -3.5 & 19.5 & 1 & 0 \\ \\ 0 & 0 & -2.3 & -2.15 & 13.55 & 0.4 & 0 \\ \hline 1 & 0 & \boxed{0.4} & 0.2 & -1.4 & -0.2 & 0 \\ 0 & 1 & -7.8 & -1.4 & 7.8 & 0.4 & 0 \end{array}$$

Vidíme, že třetí tabulka se liší od první jen rotací sloupců o dva vpravo. Použijeme-li v dalších iteracích pivotové pravidlo, které bude opět vybírat pivoty 0.4 a 2.5, tak za další čtyři iterace získáme počáteční simplexovou tabulku. ♦

Byla objevena pivotová pravidla zaručující, že algoritmus se pro žádný vstup nezacyklí. Nejznámější je *Blandovo anticyklící pravidlo*: při výběru pivotového sloupce vždy vybereme mezi sloupci s $c_j < 0$ ten s nejmenším indexem, při výběru pivotového řádku vybereme z množiny (14.9) řádek s nejmenším indexem. Důkaz správnosti pravidla vynecháme (není krátký).

14.3 Inicializace algoritmu

Zopakujme, že na začátku základního simplexového algoritmu musí být úloha zadána ve tvaru

$$\min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}, \quad (14.10)$$

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 $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ a platí $\mathbf{b} \geq \mathbf{0}$, přidáním slacků úlohu převedeme na $\min\{\mathbf{c}^T \mathbf{u} \mid \mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0}\}$. 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} jsou standardní báze.

Příklad 14.9. Vyřešte simplexovým algoritmem:

$$\begin{aligned} \min \quad & -3x_1 - x_2 - 3x_3 \\ \text{za podmíněk} \quad & 2x_1 + x_2 + x_3 \leq 2 \\ & x_1 + 2x_2 + 3x_3 \leq 5 \\ & 2x_1 + 2x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Přidáme slackové proměnné $u_1, u_2, u_3 \geq 0$, abychom omezení uvedli do tvaru rovností:

$$\begin{aligned} \min \quad & -3x_1 - x_2 - 3x_3 \\ \text{za podmíněk} \quad & 2x_1 + x_2 + x_3 + u_1 = 2 \\ & x_1 + 2x_2 + 3x_3 + u_2 = 5 \\ & 2x_1 + 2x_2 + x_3 + u_3 = 6 \\ & x_1, x_2, x_3, u_1, u_2, u_3 \geq 0 \end{aligned}$$

Zde je výchozí simplexová tabulka:

$$\begin{array}{cccccc|c} -3 & -1 & -3 & 0 & 0 & 0 & 0 \\ \hline 2 & 1 & 1 & 1 & 0 & 0 & 2 \\ 1 & 2 & 3 & 0 & 1 & 0 & 5 \\ 2 & 2 & 1 & 0 & 0 & 1 & 6 \end{array}$$

◆

14.3.1 Dvoufázová simplexová metoda

Pokud je úloha zadána v obecném tvaru, operacemi z §12.1 ji lze vždy převést do tvaru (14.10). 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 (14.10) je přípustná. V tomto případě nejdříve vyřešíme *pomocnou úlohu* LP, která najde *nějaké* (ne nutně optimální) přípustné řešení. Z něj pak získáme standardní bázi. Pomocná úloha je

$$\min\{\mathbf{1}^T \mathbf{u} \mid \mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0}\} \quad (14.11)$$

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 (14.10) je přípustná, právě když je optimální hodnota úlohy (14.11) 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 (14.10) je nepřipustná.
- Pokud je optimum rovno 0, pak úloha (14.10) je přípustná. Pokud není optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (14.11) degenerované, po skončení simplexového algoritmu jsou všechny bázové proměnné kladné. Protože $\mathbf{u} = \mathbf{0}$, proměnné \mathbf{u} budou tedy nebázové. Proto mezi sloupci příslušnými proměnným \mathbf{x} existuje standardní báze.

Pokud je optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (14.11) degenerované, některé proměnné \mathbf{u} mohou být na konci algoritmu bázové. Pak je nutno udělat dodatečné úpravy kolem pivotů ve sloupcích příslušných bázovým proměnným \mathbf{u} , abychom tyto sloupce dostali z báze ven. Toto podrobně popisovat nebudeme.

Nalezení nějakého přípustného řešení v pomocné úloze (14.11) se nazývá **první fáze** a řešení původní úlohy pak **druhá fáze** algoritmu, mluvíme tedy o **dvoufázové simplexové metodě**.

Příklad 14.10. 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 jsou standardní báze, můžeme tedy na pomocnou úlohu pustit základní simplexový algoritmus. Po vynulování ceny nad bázovými proměnnými budou kroky algoritmu vypadat takto:

$$\begin{array}{cccccc|c} -4 & -4 & -3 & 0 & 0 & 0 & -25 \\ \hline 3 & \boxed{2} & 1 & 1 & 0 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 0 & 15 \\ \hline 2 & 0 & -1 & 2 & 0 & 0 & -5 \\ \hline 1.5 & 1 & 0.5 & 0.5 & 0 & 0 & 5 \\ -2 & 0 & \boxed{1} & -1 & 1 & 0 & 5 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 2.5 & 1 & 0 & 1 & -0.5 & 0 & 2.5 \\ -2 & 0 & 1 & -1 & 1 & 0 & 5 \end{array}$$

Optimum je rovno 0, tedy původní úloha je přípustná. Proměnné u_1, u_2 jsou nebázové a tedy rovny nule, bázové proměnné jsou x_2, x_3 . Teď tedy můžeme začít druhou fází (řešení původní úlohy) s počáteční tabulkou

$$\begin{array}{ccc|c} -20 & -30 & -40 & 0 \\ \hline 2.5 & 1 & 0 & 2.5 \\ -2 & 0 & 1 & 5 \end{array}$$

◆

14.4 Cvičení

14.1. Najděte všechny báze a bázová řešení mnohostěnu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ pro

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 2 & -1 & 4 & 1 & 4 \end{array} \right].$$

Která z nich jsou přípustná? Která z nich jsou degenerovaná?

14.2. Protože $\text{rank } \mathbf{A} = m$, dimenze mnohostěnu (14.1) 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 $(n, k) \in \mathbb{N} \times \{0, \dots, n - m\}$ 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 $\dim X = k$.

14.3. V tabulce zakroužkujte všechny pivoty takové, že ekvivalentní úprava kolem nich povede k přípustnému bázovému řešení:

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|ccc} 0 & 2 & 6 & 1 & 0 & -4 & 3 & 0 & 4 \\ 1 & 1 & -3 & 0 & 0 & 2 & 3 & 0 & 3 \\ 0 & -1 & 1 & 0 & 1 & -2 & -3 & 0 & 1 \\ 0 & -2 & 2 & 0 & 0 & 2 & -1 & 1 & 1 \end{array} \right]$$

14.4. Zapište lineární program

$$\begin{array}{ll} \min & -x_1 \qquad \qquad \qquad - x_4 - 3x_5 \\ \text{za podmíněk} & 2x_1 \qquad \qquad \qquad + x_4 + x_5 + x_6 = 2 \\ & -x_1 + x_2 \qquad \qquad \qquad + 2x_4 + 3x_5 = 4 \\ & 2x_1 \qquad \qquad + x_3 + 2x_4 - x_5 = 6 \\ & \qquad \qquad \qquad \qquad \qquad \qquad x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{array}$$

do simplexové tabulky. Předpokládejte, že aktuální báze je $\{2, 3, 6\}$. Jaké je aktuální bázové řešení? Je toto bázové řešení přípustné. Je degenerované? Pokud je to možné, udělejte jeden krok simplexového algoritmu. Pokud to možné není, vysvětlete proč.

14.5. Vyřešte simplexovou metodou (nejdříve ji co nejjednodušeji inicializujte):

$$\begin{array}{ll} \max & 2x_1 - x_2 - 3x_3 \\ \text{za podmíněk} & -2x_1 - x_2 + x_3 \leq 2 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & -2x_1 - 4x_2 + x_3 \leq 6 \\ & \qquad \qquad \qquad x_1, x_2, x_3 \geq 0 \end{array}$$

14.6. Vyřešte simplexovou metodou (i když lze řešit úvahou):

$$\begin{aligned} \max \quad & 6x_1 + 9x_2 + 5x_3 + 9x_4 \\ \text{za podmínek} \quad & x_1 + x_2 + x_3 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

14.7. Necht' úloha (14.7) má více optimálních řešení. Jak se to pozná v simplexové tabulce? Navrhněte algoritmus, který vypíše všechna optimální bázevá řešení.

14.8. Mějme lineární program

$$\begin{aligned} \min \quad & 2x_1 \quad - 3x_3 + x_4 \\ \text{za podmínek} \quad & x_1 - x_2 - x_3 \geq 0 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & 2x_1 - x_2 - x_3 + 2x_4 = 6 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Inicializujte co nejjednodušším způsobem základní simplexový algoritmus. Vyřešte tímto algoritmem. Nepoužívejte dvoufázovou metodu.

14.9. Vyřešte dvoufázovou simplexovou metodou:

$$\begin{aligned} \max \quad & 3x_1 - 4x_2 \\ \text{za podmínek} \quad & -2x_1 - 5x_2 \leq 10 \\ & 3x_1 + x_2 \leq 3 \\ & -2x_1 + x_2 \leq -2 \\ & x_1 \geq 0 \\ & x_2 \leq -1 \end{aligned}$$

14.10. Necht' úloha (14.7) má simplexovou tabulku

$$\begin{bmatrix} \mathbf{c}^T & d \end{bmatrix} = \left[\begin{array}{cccccc|c} 3 & 0 & 0 & 1 & -3 & 0 & 0 \\ 7 & 0 & 0 & -3 & 3 & 2 & 1 \\ 1 & 1 & 0 & 0 & -1 & 0 & 3 \\ -1 & 0 & 2 & 1 & -1 & 0 & 1 \end{array} \right]$$

Je tato úloha omezená?

14.11. (★) Cvičení 3.22 má vztah k důkazu Důsledku 14.1. Jaký je tento vztah?

Nápověda a řešení

14.1.	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

14.2. Nápověda: mnohostěn $\{(x, y) \in \mathbb{R}^2 \mid x + y = 1, x \geq 0, y \geq 0\}$ má dimenzi 1, mnohostěn $\{(x, y) \in \mathbb{R}^2 \mid x + y = 0, x \geq 0, y \geq 0\}$ má dimenzi 0.

14.5. Úloha je neomezená kvůli prvnímu sloupci.

14.8. Např. otočíme znaménko prvního omezení, třetí omezení vydělíme dvěma, přidáme slackové proměnné pro první a druhé omezení. Pak vynulujeme ceny nad bázovými sloupci. Iterace algoritmu:

$$\begin{array}{c|cccccc|c}
 1 & 0.5 & -2.5 & 0 & 0 & 0 & -3 \\
 \hline
 -1 & 1 & \boxed{1} & 0 & 1 & 0 & 0 \\
 -1 & 2 & -3 & 0 & 0 & 1 & 5 \\
 1 & -0.5 & -0.5 & 1 & 0 & 0 & 3
 \end{array}
 \quad
 \begin{array}{c|cccccc|c}
 -1.5 & 3 & 0 & 0 & 2.5 & 0 & -3 \\
 \hline
 -1 & 1 & 1 & 0 & 1 & 0 & 0 \\
 -4 & 5 & 0 & 0 & 3 & 1 & 5 \\
 \boxed{0.5} & 0 & 0 & 1 & 0.5 & 0 & 3
 \end{array}
 \quad
 \begin{array}{c|cccccc|c}
 0 & 3 & 0 & 3 & 4 & 0 & 6 \\
 \hline
 0 & 1 & 1 & 2 & 2 & 0 & 6 \\
 0 & 5 & 0 & 8 & 7 & 1 & 29 \\
 1 & 0 & 0 & 2 & 1 & 0 & 6
 \end{array}$$

Výsledek: $(x_1, x_2, x_3, x_4) = (6, 0, 6, 0)$, hodnota optima -6 .

14.9. Optimum je $(x_1, x_2) = (25, -36)/13$.

14.10. Úloha je neomezená. Mohlo by se zdát, že je omezená, protože neexistuje sloupec j takový, že $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i . Ale po jedné či několika iteracích simplexové metody se takový sloupec objeví.

Kapitola 15

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 15.1): 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.

15.1 Konstrukce duální úlohy

K úloze LP v obecném tvaru (viz §12.1) se duální úloha získá dle tohoto postupu:

$$\begin{array}{ll} \min \sum_{j \in J} c_j x_j & \max \sum_{i \in I} b_i y_i \\ \text{za podm. } \sum_{j \in J} a_{ij} x_j = b_i & \text{za podm. } y_i \in \mathbb{R}, \quad i \in I_0 \\ & y_i \geq 0, \quad i \in I_+ \\ & y_i \leq 0, \quad i \in I_- \\ x_j \in \mathbb{R} & \sum_{i \in I} a_{ij} y_i = c_j, \quad j \in J_0 \\ x_j \geq 0 & \sum_{i \in I} a_{ij} y_i \leq c_j, \quad j \in J_+ \\ x_j \leq 0 & \sum_{i \in I} a_{ij} y_i \geq c_j, \quad j \in J_- \end{array} \quad (15.1)$$

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 15.1. Následující dvojice lineárních programů je navzájem duální:

$$\begin{array}{ll}
 \min & 2x_1 - 3x_3 + x_4 \\
 \text{za podm.} & 2x_1 - x_2 + x_3 + 2x_4 = 6 \\
 & -x_1 + 2x_2 - 3x_3 \leq 5 \\
 & x_1 - x_2 - x_3 - 3x_4 \geq 0 \\
 & x_1 \geq 0 \\
 & x_2 \in \mathbb{R} \\
 & x_3 \geq 0 \\
 & x_4 \leq 0
 \end{array}
 \qquad
 \begin{array}{ll}
 \max & 6y_1 + 5y_2 \\
 \text{za podm.} & y_1 \in \mathbb{R} \\
 & y_2 \leq 0 \\
 & y_3 \geq 0 \\
 & 2y_1 - y_2 + y_3 \leq 2 \\
 & -y_1 + 2y_2 - y_3 = 0 \\
 & y_1 - 3y_2 - y_3 \leq -3 \\
 & 2y_1 - 3y_3 \geq 1
 \end{array}
 \quad \blacklozenge$$

Pro speciální tvary LP se dvojice duálních úloh přehledněji napíše v maticové formě. Např. pro $I_0 = I_- = J_0 = J_- = \emptyset$ obdržíme

$$\begin{array}{ll}
 \min & \mathbf{c}^T \mathbf{x} \\
 \text{za podm.} & \mathbf{Ax} \geq \mathbf{b} \\
 & \mathbf{x} \geq \mathbf{0}
 \end{array}
 \qquad
 \begin{array}{ll}
 \max & \mathbf{b}^T \mathbf{y} \\
 \text{za podm.} & \mathbf{y} \geq \mathbf{0} \\
 & \mathbf{A}^T \mathbf{y} \leq \mathbf{c}
 \end{array}
 \quad (15.2)$$

15.2 Věty o dualitě

Následující věty platí pro obecný tvar (15.1), ale důkazy uděláme jen pro speciální tvar (15.2).

Věta 15.1 (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

$$\mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T \mathbf{Ax} \geq \mathbf{y}^T \mathbf{b}. \quad (15.3)$$

Uved'me jeden okamžitý důsledek slabé duality.

Důsledek 15.2. *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 15.3 (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:*

$$\sum_{j \in J} a_{ij} x_j = b_i \quad \text{nebo} \quad y_i = 0 \quad \forall i \in I, \quad (15.4a)$$

$$x_j = 0 \quad \text{nebo} \quad \sum_{i \in I} a_{ij} y_j = c_j \quad \forall j \in J. \quad (15.4b)$$

Podmínky (15.4) 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 (15.3) je jasné, že pro přípustná \mathbf{x}, \mathbf{y} platí ekvivalence

$$\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b} \iff \mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{b}. \quad (15.5)$$

Dvě rovnosti na pravé straně této ekvivalence jde napsat jako

$$\mathbf{y}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) = 0, \quad (15.6a)$$

$$(\mathbf{c}^T - \mathbf{y}^T \mathbf{A}) \mathbf{x} = 0. \quad (15.6b)$$

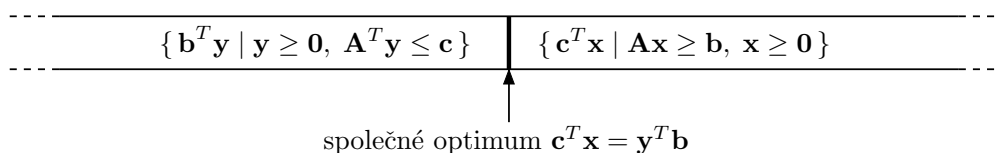
Levé strany těchto rovností jsou skalární součiny nezáporných (díky přípustnosti \mathbf{x}, \mathbf{y}) vektorů. Nyní si stačí uvědomit, že pro libovolné nezáporné vektory $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ platí

$$\mathbf{u}^T \mathbf{v} = 0 \iff \forall i (u_i v_i = 0) \iff \forall i (u_i = 0 \text{ nebo } v_i = 0). \quad \blacksquare$$

Uvědomte si, že Důsledek 15.2 a Věta 15.3 neříkají, že rovnost $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$ vůbec někdy nastane. To je předmětem nejdůležitější věty o dualitě:

Věta 15.4 (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} , pak $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$.*

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:



Příklad 15.2. Mějme dvojici navzájem duálních úloh LP:

$\begin{aligned} \min \quad & 2x_1 + 5x_2 + 6x_3 = \mathbf{5.4} \\ \mathbf{3} = \quad & 2x_1 + x_2 + 2x_3 \geq 3 \\ \mathbf{2.4} = \quad & x_1 + 2x_2 + 2x_3 \geq 1 \\ \mathbf{3} = \quad & x_1 + 3x_2 + x_3 \geq 3 \\ -\mathbf{0.6} = \quad & -x_1 + x_2 - 2x_3 \geq -1 \\ \mathbf{1.2} = \quad & x_1 \geq 0 \\ \mathbf{0.6} = \quad & x_2 \geq 0 \\ \mathbf{0} = \quad & x_3 \geq 0 \end{aligned}$	$\begin{aligned} \max \quad & 3y_1 + y_2 + 3y_3 - y_4 = \mathbf{5.4} \\ \mathbf{0.2} = \quad & y_1 \geq 0 \\ \mathbf{0} = \quad & y_2 \geq 0 \\ \mathbf{1.6} = \quad & y_3 \geq 0 \\ \mathbf{0} = \quad & y_4 \geq 0 \\ \mathbf{2} = \quad & 2y_1 + y_2 + y_3 - y_4 \leq 2 \\ \mathbf{5} = \quad & y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\ \mathbf{2} = \quad & 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6 \end{aligned}$
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Spočetli jsme optimální řešení obou úloh a dosadili tato řešení do účelových funkcí a do omezení. Hodnoty optimálních řešení $\mathbf{x}^* = (1.2, 0.6, 0)$ a $\mathbf{y}^* = (0.2, 0, 1.6)$ a hodnoty omezení a účelových funkcí v optimech jsou napsané tučně před/za rovnítky. Dle věty o silné dualitě se optima

rovnají. Vezmeme-li libovolný řádek (kromě účelového), je na něm alespoň jedno z obou omezení aktivní. Např. ve druhém řádku je primární omezení $2x_1 + x_2 + 2x_3 \geq 3$ aktivní a duální omezení $y_1 \geq 0$ je neaktivní. Podle věty o komplementaritě se nemůže stát, že by na některém řádku byly obě omezení zároveň neaktivní (mohou být obě ale zároveň aktivní, což zde nenastává, ale může to nastat v případě degenerace). ♦

Zopakujme (viz §12), že pro každou úlohu LP mohou nastat 3 možnosti: úloha má optimální řešení, úloha je neomezená, úloha je nepřipustná.

Věta 15.5. Z devíti možností pro dvojici duálních úloh se realizují tyto:

primární/duální	<i>má optimum</i>	<i>neomezená</i>	<i>nepřipustná</i>
<i>má optimum</i>	<i>ano</i>	<i>ne</i>	<i>ne</i>
<i>neomezená</i>	<i>ne</i>	<i>ne</i>	<i>ano</i>
<i>nepřipustná</i>	<i>ne</i>	<i>ano</i>	<i>ano</i>

Důkaz. Snadno najdeme příklady dvojic duálních úloh, které realizují povolené kombinace. Čtyři zakázané kombinace v prvním řádku a prvním sloupci plynou z první části věty o silné dualitě (primární úloha má optimum *právě tehdy*, když duální úloha má optimum).

Poslední zakázaný případ odůvodníme poněkud neformálně. Lze ukázat, že věta o slabé dualitě platí i tehdy, kdy jedna úloha je neomezená, přičemž pro primární [duální] neomezenou úlohu definujeme hodnotu optima (přesněji infima [suprema]) $-\infty$ [$+\infty$]. Pak tato věta zakazuje, aby úlohy byly zároveň neomezené, protože pak bychom měli $-\infty \geq +\infty$. ■

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 (tzv. **certifikát 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 15.3. 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 15.2 (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

$$\begin{aligned} 2y_1 + y_3 &= 2 \\ y_1 + 3y_3 &= 5 \end{aligned} \tag{15.7}$$

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 (15.7) 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*. ♦

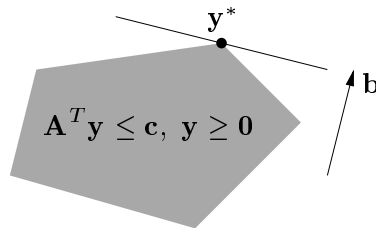
Zkoumejme, jak se změní optimální hodnota úlohy $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, jestliže nepatrně změníme pravé strany omezení \mathbf{b} . Odpověď je snadno vidět v duálu.

Věta 15.6 (o stínových cenách). Necht' funkce $f: \mathbb{R}^m \rightarrow \mathbb{R}$ je definována jako

$$f(\mathbf{b}) = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} = \max\{\mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0}\},$$

přičemž předpokládáme, že primární (a tedy i duální) úloha má optimální řešení. 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 v extrémálním bodě 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 $f'(\mathbf{b}) = \mathbf{y}^{*T}$. ■

Zdůrazněme předpoklad jednoznačnosti optimálního řešení. Kdyby množina duálních optimálních řešení byla ne jediný extrémální bod, ale stěna vyšší dimenze, po malé změně vektoru \mathbf{b} by se optimální stěna mohla stát extrémálním bodem a funkce f by tedy v bodě \mathbf{b} nebyla diferencovatelná.

Protože \mathbf{b} je zároveň vektor pravých stran primární úlohy, optimální duální proměnné \mathbf{y}^* vyjadřují *citlivost* optima primární úlohy na změnu pravých stran primárních omezení $\mathbf{A}\mathbf{x} \geq \mathbf{b}$. Interpretujeme-li naše LP jako optimální výrobní plán (12.9) (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 15.4. Necht' je známo, že duální úloha v Příkladu 15.2 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. \blacklozen

15.3 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 15.5. Demonstrujme nyní dualitu na velmi jednoduchém lineárním programu

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0} \} = \min \{ c_1 x_1 + \dots + c_n x_n \mid x_1 + \dots + x_n = 1, x_i \geq 0 \},$$

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 12.3), 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 (15.1) sestrojíme duální úlohu

$$\max \{ y \in \mathbb{R} \mid \mathbf{y} \mathbf{1} \leq \mathbf{c} \} = \max \{ y \in \mathbb{R} \mid y \leq c_i \forall i = 1, \dots, n \}.$$

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. \blacklozen

Příklad 15.6. Pro daná $\mathbf{c} \in \mathbb{R}^n$ a $k \in \{1, \dots, n\}$ máme úlohu

$$\max\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} \leq k \} \quad (15.8)$$

Duální úlohu sestrojíme dle předpisu (15.1):

$$\begin{array}{ll} \max & c_1 x_1 + \dots + c_n x_n \\ \text{za podm.} & x_1 + \dots + x_n \leq k \\ & x_i \leq 1 \\ & x_i \geq 0 \end{array} \quad \begin{array}{ll} \min & ky + z_1 + \dots + z_n \\ \text{za podm.} & y \geq 0 \\ & z_i \geq 0 \quad i = 1, \dots, n \\ & y + z_i \geq c_i \quad i = 1, \dots, n \end{array}$$

Pravá úloha odpovídá levé úloze z dvojice (15.1), protože v úloze (15.8) se maximalizuje. Sestrojit duální úlohu je zde obtížnější než např. v Příkladě 15.2 kvůli přítomnosti indexů $i = 1, \dots, n$. Doporučujeme např. napsat dvojici (15.1) 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 (15.8) se snadno vyřeší úvahou (viz Cvičení 12.3): její optimální hodnota je součet k největších kladných čísel c_i . Ovšem na první pohled není patrné, proč duální úloha má stejnou optimální hodnotu. Zkuste takovou úvahu vymyslet! \blacklozen

Příklad 15.7 (Příklad: Ekonomická interpretace duality). Vraťme se k Příkladu 1.10 o výrobci lupínků a hranolků z brambor a oleje. Napišme k této úloze duální úlohu:

$$\begin{array}{ll} \max & 120l + 76h \\ \text{za podm.} & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l \geq 0 \\ & h \geq 0 \end{array} \quad \begin{array}{ll} \min & 100a + 16b \\ \text{za podm.} & a \geq 0 \\ & b \geq 0 \\ & 2a + 0.4b \geq 120 \\ & 1.5a + 0.2b \geq 76 \end{array}$$

Přijde překupník a chce koupit od výrobce jeho zásoby brambor a oleje. Překupník řeší tuto otázku: Jaké nejnižší ceny musím nabídnout, aby mi výrobce 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 15.6), proč se optimálním duálním proměnným někdy říká *stínové ceny* odpovídajících primárních omezení. Např. stínová cena brambor je 32 Kč/kg. \blacklozen

Příklad 15.8. Z §12.4 víme, že optimální argument úlohy

$$\min_{x \in \mathbb{R}} \sum_{i=1}^n |x - a_i| = \min\{ z_1 + \dots + z_n \mid z_i \in \mathbb{R}, x \in \mathbb{R}, -z_i \leq x - a_i \leq z_i \} \quad (15.9)$$

je medián z čísel a_1, \dots, a_n . Chceme sestrojít duální úlohu a zjednodušit ji. Chceme pro tuto úlohu dokázat platnost silné duality.

Duální úlohu je možné sestrojít podle předpisu (15.1):

$$\begin{array}{ll} \min & \sum_{i=1}^n z_i \\ \text{za podm.} & x + z_i \geq a_i \\ & -x + z_i \geq -a_i \\ & z_i \in \mathbb{R} \\ & x \in \mathbb{R} \end{array} \quad \begin{array}{ll} \max & \sum_{i=1}^n (p_i - q_i)a_i \\ \text{za podm.} & p_i \geq 0 \\ & q_i \geq 0 \\ & p_i + q_i = 1 \\ & \sum_{i=1}^n (p_i - q_i) = 0 \end{array} \quad \begin{array}{l} i = 1, \dots, n \\ i = 1, \dots, n \\ i = 1, \dots, n \end{array}$$

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

$$\begin{array}{ll} \min & \mathbf{1}^T \mathbf{z} \\ \text{za podm.} & \mathbf{1}x + \mathbf{z} \geq \mathbf{a} \\ & -\mathbf{1}x + \mathbf{z} \geq -\mathbf{a} \\ & \mathbf{z} \in \mathbb{R}^n \\ & x \in \mathbb{R} \end{array} \quad \begin{array}{ll} \max & \mathbf{a}^T (\mathbf{p} - \mathbf{q}) \\ \text{za podm.} & \mathbf{p} \geq \mathbf{0} \\ & \mathbf{q} \geq \mathbf{0} \\ & \mathbf{p} + \mathbf{q} = \mathbf{1} \\ & \mathbf{1}^T (\mathbf{p} - \mathbf{q}) = 0 \end{array}$$

neboli

$$\begin{array}{ll} \min & \mathbf{h}^T \mathbf{u} \\ \text{za podm.} & \mathbf{F}\mathbf{u} \geq \mathbf{g} \\ & \mathbf{u} \in \mathbb{R}^{1+n} \end{array} \quad \begin{array}{ll} \max & \mathbf{g}^T \mathbf{v} \\ \text{za podm.} & \mathbf{v} \geq \mathbf{0} \\ & \mathbf{F}^T \mathbf{v} = \mathbf{h} \end{array}$$

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

$$\max \{ a_1 t_1 + \dots + a_n t_n \mid t_i \in \mathbb{R}, t_1 + \dots + t_n = 0, -1 \leq t_i \leq 1 \}. \quad (15.10)$$

Primární úloha (15.9) a duální úloha (15.10) 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 (15.9) 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 (15.10), 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 $\sum_i a_i t_i = \sum_i |a_i|$. ♦

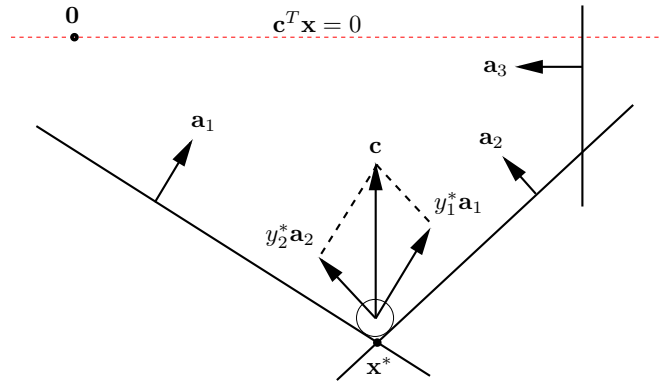
15.3.1 Mechanické modely

Lineární programy lze modelovat mechanickými (obecněji fyzikálními) modely. Můžeme jim říkat mechanické či analogové počítače. Takové modely se dobře hodí na ilustraci duality.

Příklad 15.9. Uvažujme dvojici duálních úloh

$$\min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n \} = \max\{ \mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0} \}. \quad (15.11)$$

Mějme mnohostěn tvořený poloprostory $\mathbf{a}_i^T \mathbf{x} \geq b_i$ a vektor \mathbf{c} mířící svisle vzhůru:



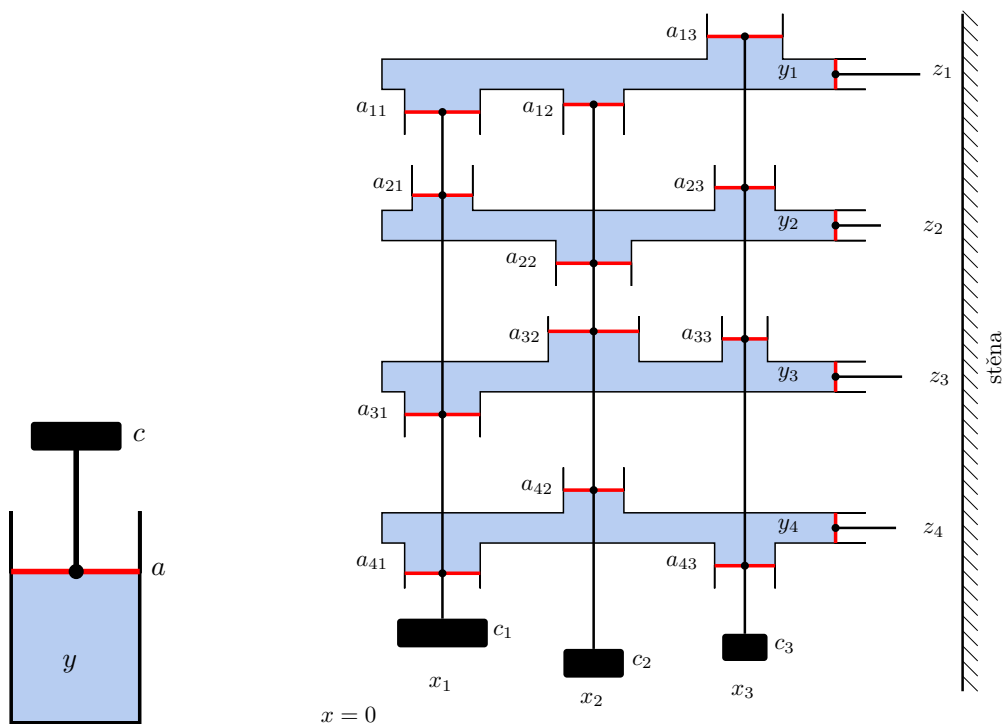
Vhodíme do mnohostěnu maličký míček, na který působí tíhová síla $-\mathbf{c}$. Pro míček se středem v bodě \mathbf{x} je číslo $\mathbf{c}^T \mathbf{x}$ přímo úměrné výšce míčku nad vodorovnou rovinou procházející počátkem, tedy potenciální energii míčku. Míček se zastaví v místě s nejmenší potenciální 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 síly 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 15.3). 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 $\mathbf{b}^T \mathbf{y}^*$. Odstraňme 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 $\sum_i b_i y_i^* = \mathbf{b}^T \mathbf{y}^*$. ♦

Příklad 15.10. Známý fyzikální zákon říká, že v nádobě s nestačitelnou 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 i -tou tyčí a stěnou je z_i , přičemž při $\mathbf{x} = \mathbf{0}$ je $\mathbf{z} = -\mathbf{b}$. Tíha j -tého závaží je c_j . Tlak v i -té nádrži je y_i .

Ukážeme, že stroj 'řeší' levou (primární) úlohu z dvojice (15.11). Ze zachování objemu kapaliny v nádobě i plyne $z_i = a_{i1}x_1 + \dots + a_{in}x_n - b_i$, tedy $\mathbf{z} = \mathbf{A}\mathbf{x} - \mathbf{b}$. Protože vodorovné tyče nemohou projít stěnou, musí vždy být $z_i \geq 0$, tedy $\mathbf{z} \geq \mathbf{0}$. 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 $y \geq 0$. 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.

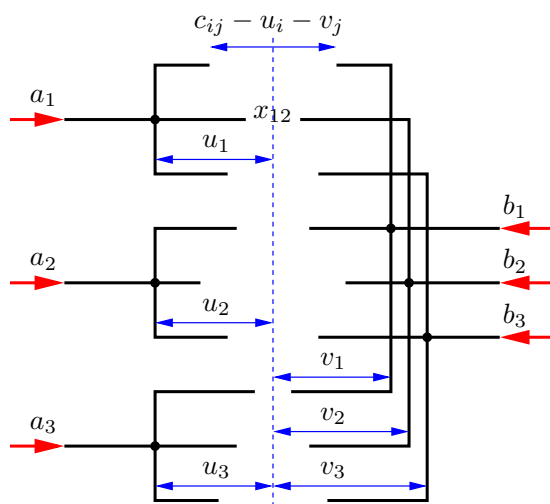
Věta o stínových cenách říká, že se změnou b_i se hodnota minimální potenciální energie mění tím více, čím je větší tlak y_i . To je ale jasné, protože čím je větší tlak, tím větší práce je třeba na posunutí tyče od stěny do vzdálenosti b_i . ♦

Mohlo by se zdát, že mechanické modely z předchozích dvou příkladů dokazují větu o silné dualitě pro dvojici úloh (15.11). 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 15.11. Zkoumejme duální úlohu k dopravní úloze (12.11). Konstrukci duálu už nebudeme popisovat podrobně, výsledná dvojice navzájem duálních úloh je

$$\begin{array}{ll} \min & \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \\ \text{za podm.} & \sum_{j=1}^n x_{ij} = a_i \\ & \sum_{i=1}^m x_{ij} = b_j \\ & x_{ij} \geq 0 \end{array} \quad \begin{array}{ll} \max & \sum_{i=1}^m a_i u_i + \sum_{j=1}^n b_j v_j \\ \text{za podm.} & u_i \in \mathbb{R} \\ & v_j \in \mathbb{R} \\ & u_i + v_j \leq c_{ij} \end{array} \quad \begin{array}{l} i = 1, \dots, m \\ j = 1, \dots, n \\ i = 1, \dots, m, j = 1, \dots, n \end{array}$$

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! ♦

15.4 Cvičení

15.1. Dokažte, že duál duálu se rovná původní úloze. Udělejte pro (a) dvojici (15.2), (b) dvojici (15.1).

15.2. Pro daná čísla $c_1, \dots, c_n \in \mathbb{R}$ chceme maximalizovat $\sum_{i=1}^n c_i x_i$ za podmínek $-1 \leq x_i \leq 1$.

²Chování stroje samozřejmě není polohou této referenční roviny ovlivněno. Jak se to projevuje algebraicky v primární a duální úloze?

- a) Vyřešte úvahou.
 b) Sestrojte duální úlohu a upravte ji do co nejjednoduššího tvaru. Vyřešte duální úlohu úvahou (musí vám vyjít stejná optimální hodnota jako u primární úlohy).
 c) Napište podmínky komplementarity.
 d) Najděte číselné hodnoty optimálních primárních a duálních proměnných (které si odpovídají přes podmínky komplementarity) pro $n = 3$ a $(c_1, c_2, c_3) = (-2, 3, 4)$.

15.3. Napište duální úlohu a podmínky komplementarity k následujícím úlohám. Pokud úloha není LP, nejdříve převed'te na LP (dle §12.1). 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 15.5.

- a) lineární program ze Cvičení 14.8
 b) $\min_{x \in \mathbb{R}} \max_{i=1}^n |a_i - x|$ (střed intervalu)
 c) úloha (12.16) (přibližné řešení přeuročené lin. soustavy ve smyslu max-normy)
 d) úloha (12.18) (přibližné řešení přeuročené lin. soustavy ve smyslu 1-normy)
 e) všechny úlohy ze Cvičení 12.3
 f) úloha vzniklá ve Cvičení 12.10 o kladce se závažími
 g) minimalizace maxima afinních funkcí (viz §12.1.1):

- (i) $\min_{x_1, x_2 \in \mathbb{R}} \max\{2x_1 - x_2 - 3, 1 - x_1, x_2 - 2, x_1 + x_2\}$
 (ii) $\min_{\mathbf{x} \in \mathbb{R}^n} \max_{i=1}^m (\mathbf{a}_i^T \mathbf{x} + b_i)$

15.4. Dokažte bez užití algoritmu na řešení LP, že $\mathbf{x} = (1, 1, 1, 1)$ je optimální řešení úlohy

$$\begin{aligned} & \min \quad [47 \quad 93 \quad 17 \quad -93] \mathbf{x} \\ & \text{za podm.} \quad \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix} \end{aligned}$$

Nápověda a řešení

15.1. Pravou úlohu nejdříve převedeme na tvar levé úlohy a pak k ní napíšeme duální úlohu. Ukážeme jen pro (15.2):

$$\begin{aligned} & - \min \quad (-\mathbf{b})^T \mathbf{y} & & - \max \quad (-\mathbf{c})^T \mathbf{x} \\ & \text{za podm.} \quad \begin{matrix} (-\mathbf{A}^T) \mathbf{y} \geq -\mathbf{c} \\ \mathbf{y} \geq \mathbf{0} \end{matrix} & & \text{za podm.} \quad \begin{matrix} \mathbf{x} \geq \mathbf{0} \\ (-\mathbf{A}) \mathbf{x} \leq -\mathbf{b} \end{matrix} \end{aligned} \quad (15.12)$$

Levá úloha (15.12) je ekvivalentní pravé úloze (15.2), pravá úloha (15.12) je ekvivalentní levé úloze (15.2).

15.2.a) Optimální hodnota je $\sum_{i=1}^n |c_i|$ (viz Cvičení 12.3)

15.2.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 $\sum_i |c_i|$.

15.2.c) Pro každé i platí $x_i = -1$ nebo $u_i = 0$. Pro každé i platí $x_i = 1$ nebo $v_i = 0$.

15.2.d) $(x_1, x_2, x_3) = (-1, 1, 1)$, $(u_1, u_2, u_3) = (2, 0, 0)$, $(v_1, v_2, v_3) = (0, 3, 4)$.

15.3.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: $z_i(y_i - m_i) = 0$, $z'_i(y'_i - m'_i) = 0$, $(z_i - d_i - x)y_i = 0$, $(z'_i - d'_i + x)y'_i = 0$.

15.3.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. $2u_1 - u_2 + u_4 = 0$, $-u_1 + u_3 + u_4 = 0$, $u_1 + \dots + u_4 = 1$, $u_1, \dots, u_4 \geq 0$.

15.4. Postupujte podle Příkladu 15.3.

Část IV
Konvexní optimalizace

Kapitola 16

Konvexní funkce

Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je **konvexní** na konvexní množině $X \subseteq \mathbb{R}^n$, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}). \quad (16.1)$$

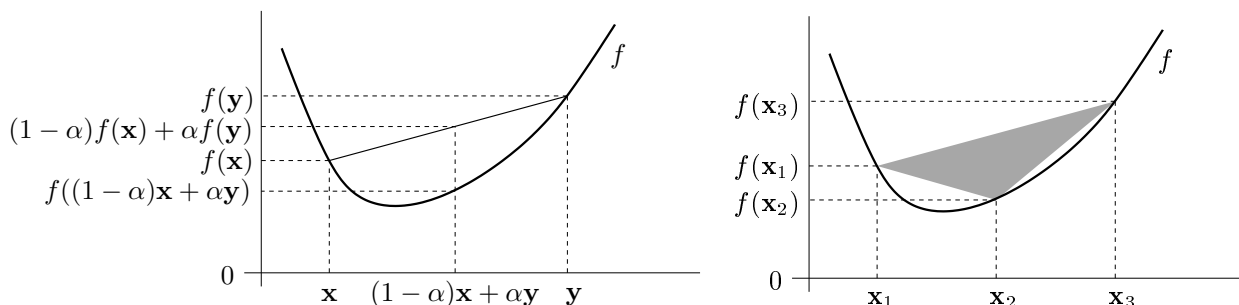
Funkce f je **konkávní** na množině X , jestliže je funkce $-f$ konvexní na množině X . Rozlišujte pojmy *konvexní množina* a *konvexní funkce*, jde o různé věci. Všimněte si, že množina X musí být konvexní (pojem konvexní nebo konkávní funkce na nekonvexní množině není definován). Pokud X je celý definiční obor funkce f , odkaz na X můžeme vynechat a říkáme jen, že funkce f je konvexní.

Podmínku (16.1) lze zobecnit pro více než dva body: funkce f je konvexní, právě když

$$\left. \begin{array}{l} \mathbf{x}_1, \dots, \mathbf{x}_k \in X \\ \alpha_1, \dots, \alpha_k \geq 0 \\ \alpha_1 + \dots + \alpha_k = 1 \end{array} \right\} \implies f(\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k) \leq \alpha_1 f(\mathbf{x}_1) + \dots + \alpha_k f(\mathbf{x}_k). \quad (16.2)$$

Podmínka (16.2) je známá jako **Jensenova nerovnost**. Podmínka (16.1) je očividně speciální případ podmínky (16.2). Naopak lze dokázat, že (16.1) implikuje (16.2). Porovnejte s definicí lineárního zobrazení (3.8) a afinního zobrazení!

Geometrický význam podmínky (16.1) 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 (16.2) je ten, že konvexní mnohoúhelník vybarvený šedě (viz pravý obrázek) leží nad grafem funkce. Podrobně rozmyslete, jak tyto geometrické interpretace odpovídají výrazům (16.1) a (16.2)!



Příklad 16.1. Dokažme z podmínky (16.1), 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

$$f((1 - \alpha)\mathbf{x} + \alpha\mathbf{y}) = \max_i((1 - \alpha)x_i + \alpha y_i) \quad (16.3a)$$

$$\leq \max_i(1 - \alpha)x_i + \max_i \alpha y_i \quad (16.3b)$$

$$= (1 - \alpha) \max_i x_i + \alpha \max_i y_i \quad (16.3c)$$

$$= (1 - \alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}) \quad (16.3d)$$

kde rovnost (16.3c) plyne z nezápornosti čísel α a $1 - \alpha$.

Nerovnost (16.3b) plyne z toho, že pro každé $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$ platí

$$\max_i(a_i + b_i) \leq \max_i a_i + \max_i b_i. \quad (16.4)$$

Nerovnost (16.4) 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 $\max_i(a_i + b_i) = a_{i^*} + b_{i^*} \leq a_{j^*} + b_{k^*} = \max_i a_i + \max_i b_i$. \blacklozen

Příklad 16.2. Dokažme z podmínky (16.1), 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 (16.1), neboť

$$f((\mathbf{x} + \mathbf{y})/2) = f(1, 1) = 1 > (f(\mathbf{x}) + f(\mathbf{y}))/2 = (0 + 0)/2 = 0. \quad \blacklozen$$

Použitím Jensenovy nerovnosti na vhodnou konvexní funkci lze získat mnoho užitečných nerovností.

Příklad 16.3. 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 (16.2) (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. \blacklozen

Příklad 16.4. 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 16.5).
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 §12.4.1).

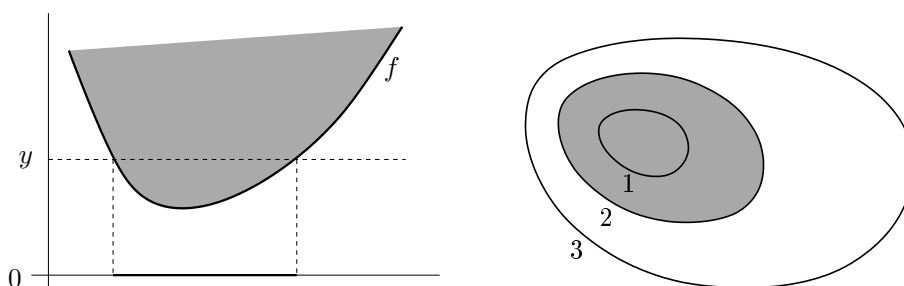
Načrtněte vrstevnice a grafy těchto funkcí (v případě více proměnných jen pro $n \in \{1, 2\}$)! ♦

16.1 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 16.1. *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

$$f((1 - \alpha)\mathbf{x}_1 + \alpha\mathbf{x}_2) \leq (1 - \alpha)f(\mathbf{x}_1) + \alpha f(\mathbf{x}_2) \leq (1 - \alpha)y_1 + \alpha y_2,$$

¹Slovo 'subkontura' je pokus o český překlad anglického 'sublevel set'.

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. ■

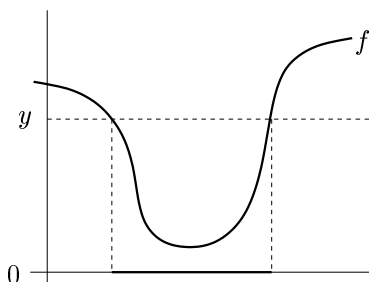
Věta 16.2. Každá subkontura konvexní funkce je konvexní množina.

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 (13.1) je tedy subkontura konvexní množina. ■

Obrácená implikace ve Větě 16.2 neplatí: existuje funkce, která není konvexní a jejíž každá subkontura je konvexní množina². Např. každá subkontura monotonní (tj. nerostoucí nebo neklesající) funkce jedné proměnné je interval, tedy konvexní množina. Jiný příklad je na obrázku:



16.2 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 (16.1) (které se někdy říká podmínka nultého řádu) charakterizovat pomocí derivací. Následující dvě věty uvedeme bez důkazů.

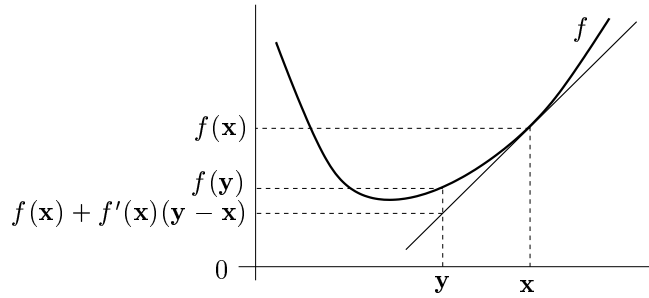
Věta 16.3 (Podmínka prvního řádu). Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná. Funkce f je konvexní (na celém \mathbb{R}^n)^a, právě když

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \implies f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

^aJde věta zobecnit pro případ, kdy ověřujeme konvexitu funkce f na konvexní podmnožině $X \subseteq \mathbb{R}^n$? Ano, ale potřebujeme navíc předpoklad, že množina X je otevřená, tj. neobsahuje žádný svůj hraniční bod. Podobně pro Větu 16.4.

²Funkce, jejíž každá subkontura je konvexní množina, se nazývá kvazikonvexní. Kvazikonvexní funkce jsou užitečné, ale zdaleka ne tak jako konvexní funkce.

To znamená, že Taylorův polynom prvního řádu funkce f v každém bodě $\mathbf{x} \in X$ (viz (8.23b)) je všude (tj. pro každé \mathbf{y}) menší nebo roven funkci f :



Věta 16.4 (Podmínka druhého řádu). Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je dvakrát diferencovatelná. Funkce f je konvexní (na celém \mathbb{R}^n), právě když v každém bodě $\mathbf{x} \in \mathbb{R}^n$ je Hessova matice $f''(\mathbf{x})$ pozitivně semidefinitní.

Příklad 16.5. 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 16.4. To je triviální, protože Hessián je $f''(\mathbf{x}) = 2\mathbf{A}$ a tedy je pozitivně semidefinitní.
- Dokažme konvexitu z Věty 16.3. 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

$$\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y} = (\mathbf{x} - \mathbf{y})^T \mathbf{A} (\mathbf{x} - \mathbf{y}), \quad (16.5)$$

což je nezáporné pro každé \mathbf{x}, \mathbf{y} , protože \mathbf{A} je pozitivně semidefinitní.

- Dokážme konvexitu z podmínky (16.1). Musíme dokázat, že pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $0 \leq \alpha \leq 1$ platí (16.1), 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 (16.5) jsme již ukázali. \blacklozenge

16.3 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 (16.1) nebo Vět 16.3 a 16.4. Dále uvedeme příklady takových operací.

³Všimněte si, že pro $n = 1$ a $\mathbf{A} = 1$ se rovnost (16.5) zjednoduší na známé $x^2 - 2xy + y^2 = (x - y)^2$.

16.3.1 Nezáporná lineární kombinace

Věta 16.5. 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

$$f = \alpha_1 g_1 + \dots + \alpha_k g_k \quad (16.6)$$

je konvexní.

Důkaz je snadný z podmínky (16.1) (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 konvexní a $f + g$ konvexní je. Např. funkce x^3 není konvexní, ale funkce $x^3 - x^3$ (tedy konstantní nulová funkce) konvexní je.

16.3.2 Skládání funkcí

Složení konvexních funkcí nemusí být konvexní funkce.

Příklad 16.6. 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!). ♦

Věta 16.6. 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. ■

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 16.6 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 16.7. 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

$$\begin{aligned} h((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) &= g(\mathbf{A}[(1-\alpha)\mathbf{x} + \alpha\mathbf{y}] + \mathbf{b}) \\ &= g((1-\alpha)(\mathbf{A}\mathbf{x} + \mathbf{b}) + \alpha(\mathbf{A}\mathbf{y} + \mathbf{b})) \\ &\leq (1-\alpha)g(\mathbf{A}\mathbf{x} + \mathbf{b}) + \alpha g(\mathbf{A}\mathbf{y} + \mathbf{b}) \\ &= (1-\alpha)h(\mathbf{x}) + \alpha h(\mathbf{y}). \end{aligned} \quad \blacksquare$$

Příklad 16.7. Mějme funkci $f: \mathbb{R}^{2n} \rightarrow \mathbb{R}$ danou vzorcem $f(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, kde $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ je libovolná norma. Tato funkce je konvexní funkce argumentu $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2n}$. Ve Větě 16.7 vezmeme $\mathbf{A} = [\mathbf{I} \quad -\mathbf{I}] \in \mathbb{R}^{n \times (2n)}$ a $\mathbf{b} = \mathbf{0}$. ♦

16.3.3 Maximum

Nejzajímavější operace zachovávající konvexitu funkcí je ovšem maximum.

Věta 16.8. *Necht' I je libovolná množina a $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I$, jsou konvexní funkce. Pak*

$$f(\mathbf{x}) = \max_{i \in I} g_i(\mathbf{x}) \quad (16.7)$$

je konvexní funkce, kde předpokládáme, že pro každé \mathbf{x} maximum existuje^a.

^aPokud 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 (16.7) nahradit supremem a věta stále platí.

Důkaz. Epigraf funkce f průnik epigrafů funkcí g_i , neboť

$$\begin{aligned} \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, \max_{i \in I} g_i(\mathbf{x}) \leq y\} &= \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, g_i(\mathbf{x}) \leq y \forall i \in I\} \\ &= \bigcap_{i \in I} \{(\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, g_i(\mathbf{x}) \leq y\}, \end{aligned}$$

kde jsme využili ekvivalence (12.7). Protože funkce g_i jsou konvexní, dle Věty 16.1 jsou jejich epigrafy konvexní množiny. Dle Věty 13.1 je průnik konvexních množin konvexní množina. Tedy epigraf funkce (16.7) je konvexní množina. Dle Věty 16.1 je tedy funkce f konvexní. ■

Příklad 16.8. Konvexitu funkce $f(\mathbf{x}) = \max_{i=1}^n x_i$ jsme již v Příkladu 16.1 dokázali z podmínky (16.1). Ovšem je mnohem pohodlnější použít Větu 16.8. 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. ◆

Příklad 16.9. 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 §12.1.1. Protože afinní funkce jsou konvexní, je i jejich maximum konvexní. ◆

Příklad 16.10. 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 nejbližšího bodu množiny C (zde předpokládáme, že maximum existuje). Dle Věty 16.7 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. ◆

Příklad 16.11. 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 §12). 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. ◆

Příklad 16.12. 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 minimem lineárních funkcí. \blacklozenge

16.4 Cvičení

16.1. Pro každou funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ dokažte z podmínky (16.1), které z těchto čtyř tvrzení platí (a pro jaké n): funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní.

- $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
- $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$
- $f(\mathbf{x}) =$ aritmetický průměr čísel x_1, \dots, x_n
- $f(\mathbf{x}) = \text{median}_{i=1}^n x_i$ (medián čísel x_1, \dots, x_n)
- $f(\mathbf{x}) = \min_{i=1}^n |x_i|$
- $f(\mathbf{x}) =$ součet dvou nejmenších čísel z čísel x_1, \dots, x_n .

16.2. Dokažte konvexitu či konkavitu funkcí z Příkladu 16.4. Můžete použít podmínku (16.1) a věty z §16.2 a §16.3.

16.3. Pro každou funkci dokažte, které z těchto čtyřech tvrzení platí: funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní. Můžete použít podmínku (16.1) a věty z §16.2 a §16.3.

- $f(x) = e^{x^2}$
- $f(x) = e^{-x^2}$
- $f(x, y) = |x - y|$
- $f(x, y) = -y$
- $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$
- $f(\mathbf{x}) = \sum_{i=1}^n x_i \log x_i$ na množině \mathbb{R}_{++}^n
- $f(\mathbf{x}) = \sum_{i=1}^k \log(b_i - \mathbf{a}_i^T \mathbf{x})$ na množině $X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} < b_i, i = 1, \dots, k\}$
- $f(\mathbf{x}) = \max_{i=1}^n x_i + \min_{i=1}^n x_i$
- $f(\mathbf{x}) = \max_{i=1}^n x_i - \min_{i=1}^n x_i$
- $f(\mathbf{x}) = \max \left\{ \left| \sum_{i=1}^n x_i - 1 \right|, \sum_{i=1}^n |x_i| - 1 \right\}$
- (*) $f(\mathbf{x}) =$ součet k největších čísel x_1, \dots, x_n (kde $k \leq n$ je dáno)

16.4. 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.

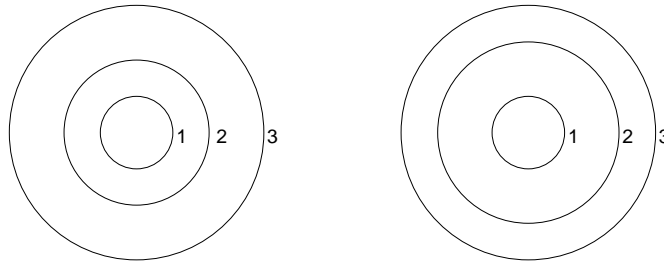
16.5. Může být součet nekonvexních funkcí konvexní funkce? Najděte protipříklad. Je to v rozporu c Větou 16.5?

16.6. 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.

16.7. 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 (16.1).



16.8. Co je subkontura výšky 2 funkce jedné proměnné $f(x) = x^2 - x$?

16.9. 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 §6.3.2) je konvexní množina.

16.10. 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).

16.11. Dokažte, že následující funkce 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 (17.3)?

b) Příklad 10.11

c) Cvičení 10.7

Nápověda a řešení

16.1.a) Konvexní i konkávní, nerovnost (16.1) platí s rovností.

16.1.b) Je konvexní, není konkávní.

16.1.c) Konvexní i konkávní, nerovnost (16.1) platí s rovností.

16.1.d) Pro $n \leq 2$ konvexní i konkávní, pro $n > 2$ ani konvexní ani konkávní.

16.1.e) Pro $n = 1$ je funkce konvexní. Pro $n = 2$ není konvexní, neboť (16.1) 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 (16.1) volbou $\mathbf{x} = -1$, $\mathbf{y} = 1$, $\alpha = \frac{1}{2}$. Pro $n > 1$ můžeme vektory opět doplnit opakováním poslední číslice.

16.1.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 $f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) = f(2, 2, 2) = 4 \not\leq (3 + 3)/2 = \frac{1}{2}f(1, 2, 3) + \frac{1}{2}f(3, 2, 1) = \frac{1}{2}f(\mathbf{x}) + \frac{1}{2}f(\mathbf{y})$.

16.3.j) Je konvexní. Absolutní hodnota je konvexní funkce, jejich součet také, maximum konvexních funkcí je konvexní funkce.

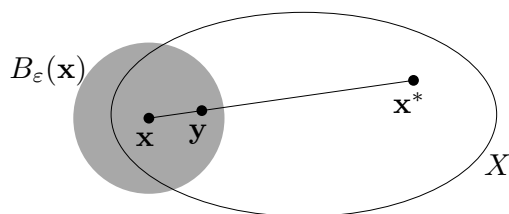
- 16.7. V podmínice (16.1) zvolte \mathbf{x}, \mathbf{y} na vrstevnicích výšky 1 a 3. Zvolte chytré α . Odpovědi: ne, ano.
- 16.8. Interval $[-1, 2]$.
- 16.9. Funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je konvexní, viz Příklad 16.5. Elipsoid je subkontura této funkce, tedy je to konvexní množina.
- 16.10. Zlogaritmováním máme $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \prod_{i=1}^n x_i \geq 1\} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_{i=1}^n \log x_i \geq 0\}$ (protože logaritmus je rostoucí funkce a tedy nemění znaménko nerovnosti). Funkce $f(\mathbf{x}) = -\sum_i \log x_i$ je na množině \mathbb{R}_{++}^n konvexní a proto její subkontura je konvexní množina.

Kapitola 17

Konvexní optimalizační úlohy

Věta 17.1. 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 §9.3) tedy existuje $\varepsilon > 0$ tak, že $f(\mathbf{x}) \leq f(\mathbf{y})$ pro všechna $\mathbf{y} \in B_\varepsilon(\mathbf{x}) \cap X$. Necht' ale \mathbf{x} není globální minimum, tedy existuje $\mathbf{x}^* \in X$ tak, 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 $B_\varepsilon(\mathbf{x})$. 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. ■

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.

17.1 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 17.1. 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. ◆

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 17.2. Uved'me příklad, na kterém bude na první pohled vidět, že nekonvexní úloha může mít velmi mnoho lokálních minim. Řešme úlohu

$$\max\{\mathbf{x}^T \mathbf{x} \mid \mathbf{x} \in [-1, 1]^n\}, \quad (17.1)$$

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

$$\min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in [-1, 1]^n\}, \quad (17.2)$$

kde $\mathbf{A} \in \mathbb{Z}^{n \times n}$ (tedy matice má celočíselné prvky). Je známo, že vyřešení (tj. nalezení globálního maxima) této úlohy pro větší n je prakticky nemožné (přesněji, je NP-těžké, srov. §12.5). ♦

Příklad 17.3. Podívejme se znovu na úlohu shlukování, Příklad 1.15 z úvodní kapitoly. Tam se minimalizuje funkce

$$f(\mathbf{x}) = f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^m \min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\| \quad (17.3)$$

přes vektory $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$. Máme $f: \mathbb{R}^{dn} \rightarrow \mathbb{R}$, tedy vlastně minimalizujeme přes jediný vektor $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{dn}$.

Je funkce (17.3) konvexní? Pro každé i je $\|\mathbf{a}_i - \mathbf{x}_j\|$ funkce vektoru \mathbf{x}_j , tedy i vektoru \mathbf{x} . Ale funkce $\min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\|$ již konvexní být nemusí (Větu 16.8 nelze použít, ta hovoří o *maximu* konvexních funkcí). Tedy ani funkce f , která je jejich součtem, nemusí být konvexní.

Tím, že se nám nepodařilo dokázat konvexitu funkce f , jsme samozřejmě nedokázali její nekonvexitu. To lze udělat následovně. Vezměme jednoduchý případ $d = 1, m = 1, n = 2, a_1 = 0$. Pak (1.23) má tvar $f(x_1, x_2) = \min\{|x_1|, |x_2|\}$. Tato funkce není konvexní, protože např. její řez $\varphi(t) = f(t, t - 1) = \min\{|t|, |t - 1|\}$ není konvexní (nakreslete si graf funkce φ). Bez důkazu uved'me, že funkce není konvexní ani pro větší d, m, n . To nás nepřekvapí, protože, jak jsme řekli Příkladu 1.15, že minimalizace funkce (17.3) je NP-těžké. ♦

Příklad 17.4. Úloha celočíselného programování (12.23) 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 13.1.g). ♦

17.2 Konvexní optimalizační úloha ve standardním tvaru

Uvažujme nyní obecnou úlohu spojitě optimalizace ve standardním tvaru (1.10),

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \quad (17.4)$$

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 17.5. Platí (promysleť!)

$$\{(x, y) \in \mathbb{R}^2 \mid x/(1+y^2) \leq 0, (x+y)^2 = 0\} = \{(x, y) \in \mathbb{R}^2 \mid x \leq 0, x+y=0\}.$$

Tedy máme dva různé popisy stejné množiny. V prvním tvaru funkce $g(x, y) = x/(1+y^2)$ není konvexní a funkce $h(x, y) = (x+y)^2$ není afinní. Přesto je množina konvexní, což je vidět ze druhého tvaru. \blacklozen

Úloze tvaru (17.4), ve které jsou funkce f, g_1, \dots, g_m konvexní a zobrazení \mathbf{h} afinní, říkáme **konvexní optimalizační úloha ve standardním tvaru**.

17.3 Ekvivalentní transformace úlohy

Dvě úlohy ve tvaru (17.4) nazveme **ekvivalentní**, když se z množiny optimálních řešení jedné dá 'snadno' získat množina optimálních řešení druhé a naopak. **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 (17.4) 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 §16.3.2).

- *Monotónní transformace účelové funkce.* Necht' $\psi: \mathbb{R} \rightarrow \mathbb{R}$ je rostoucí funkce. Pak

$$\operatorname{argmin}\{f(\mathbf{x}) \mid \mathbf{x} \in X\} = \operatorname{argmin}\{\psi(f(\mathbf{x})) \mid \mathbf{x} \in X\}.$$

Tato transformace nemusí zachovat konvexitu funkce f .

Příklad 17.6. 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. \blacklozen

- *Slackové proměnné.* Podobně jako v LP (viz §12.1), úloha (17.4) 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.* Platí (už jsme to viděli v (12.5))

$$\min\{f(\mathbf{x}) \mid \mathbf{x} \in X\} = \min\{y \mid \mathbf{x} \in X, y \in \mathbb{R}, f(\mathbf{x}) - y \leq 0\}.$$

V nové úloze vlastně hledáme nejnižší bod epigrafu funkce f na množině přípustných řešení. Plyne z toho, že každou optimalizační úlohu lze převést na úlohu s lineární účelovou funkcí. Tato transformace zachovává konvexitu úlohy: je-li X konvexní množina a f konvexní funkce na X , pak funkce $f(\mathbf{x}) - y$ proměnných (\mathbf{x}, y) je konvexní na množině $X \times \mathbb{R}$.

17.4 Třídy konvexních optimalizačních úloh

Optimalizační úlohy ve tvaru (17.4) 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>.

17.4.1 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 12, že již tento jednoduchý případ má velmi mnoho aplikací.

17.4.2 Kvadratické programování (QP)

V *kvadratickém programování* jsou funkce g_i, h_i afinní a funkce f je kvadratická. Tedy je to úloha

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} \\ \text{za podm.} \quad & \mathbf{C} \mathbf{x} \leq \mathbf{d}, \\ & \mathbf{E} \mathbf{x} = \mathbf{f}, \end{aligned}$$

kde $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{C} \in \mathbb{R}^{m \times n}$, $\mathbf{d} \in \mathbb{R}^m$, $\mathbf{E} \in \mathbb{R}^{l \times n}$, $\mathbf{f} \in \mathbb{R}^l$. Úloha je to konvexní, právě když funkce f je navíc konvexní, neboli matice \mathbf{A} je pozitivně semidefinitní (viz Příklad 16.5).

Příklad 17.7. Přibližné řešení přeурčené lineární soustavy ve smyslu nejmenších čtverců (5.2) je konvexní úloha QP bez omezení, tj. f je kvadratická konvexní (jak jsme ukázali v Příkladě 6.9) a $m = l = 0$. Tuto úlohu lze převést na řešení soustavy lineárních rovnic (5.3). ♦

Příklad 17.8. Řešení lineární soustavy s nejmenší normou (§5.2) nebo, obecněji, úloha nejmenších čtverců s omezeními typu rovnosti (Příklad 11.4) jsou příklady konvexního QP s omezeními typu rovnosti (tj. \mathbf{A} je pozitivně semidefinitní, $m = 0$ a h_i jsou afinní). Tato úloha jde převést na řešení lineární soustavy. ♦

Příklad 17.9. Obecněji, minimalizace kvadratické funkce za podmínek typu rovnosti (tj. $m = 0$, $l > 0$ a h_i jsou afinní) se dá převést na řešení soustavy lineárních rovnic, viz Cvičení 11.19. Je-li ovšem \mathbf{A} indefinitní, musíme ověřit podmínky druhého řádu (§11.2.3). ♦

Příklad 17.10. Často je užitečné úlohu nejmenších čtverců (5.2) řešit za omezení $\mathbf{c} \leq \mathbf{x} \leq \mathbf{d}$, tj. každá proměnná x_j musí být v intervalu $[c_j, d_j]$ (tzv. *box constraints*). To vede na konvexní QP s omezeními typu nerovnosti (tj. \mathbf{A} je pozitivně semidefinitní a $m > 0$). Tuto úlohu již nelze převést na řešení soustavy lineárních rovnic. ♦

Příklad 17.11. Čtverec vzdálenosti 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

$$\min\{\|\mathbf{y} - \mathbf{x}\|_2^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \leq \mathbf{b}\}. \quad \blacklozenge$$

Příklad 17.12. 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

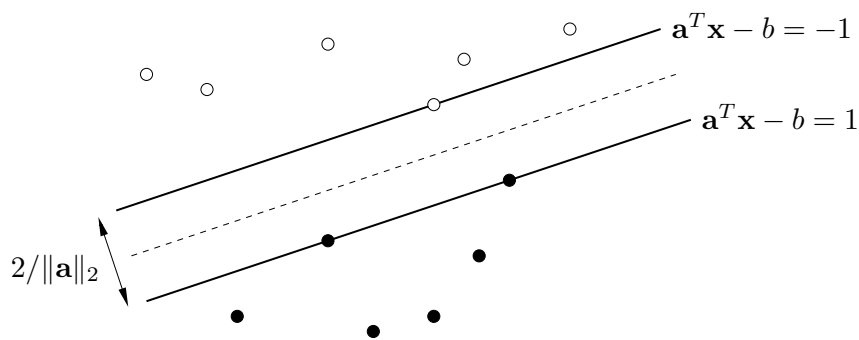
$$y_i(\mathbf{a}^T \mathbf{x}_i - b) > 0, \quad i = 1, \dots, m. \quad (17.5)$$

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 (17.5) můžeme ekvivalentně psát jako

$$y_i(\mathbf{a}^T \mathbf{x}_i - b) \geq 1, \quad i = 1, \dots, m. \quad (17.6)$$

Hledáme-li libovolnou oddělovací nadrovinu, stačí nám najít libovolné řešení soustavy nerovnic (17.6), 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 oddělovací nadrovinu která maximalizuje šířku pásu, tedy minimalizuje $\|\mathbf{a}\|_2^2 = \mathbf{a}^T \mathbf{a}$ za podmínek (17.6). To je konvexní úloha QP. ♦

17.4.3 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 a funkce h_i jsou afinní. Úloha je konvexní, právě když funkce f, g_i jsou navíc konvexní.

17.4.4 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

$$g_i(\mathbf{x}) = \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^T\mathbf{x} + d_i). \quad (17.7)$$

Tedy úloha SOCP má tvar (pro jednoduchost neuvažujeme 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, viz Příklad 16.4 a dále viz Věta 16.7). 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 eukleidovské normy $\|\cdot\|_2$, 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 17.13. Podívejme se znovu na úlohu nalzení geometrického mediánu, Příklad 1.14 z úvodní kapitoly. Tam minimalizujeme funkci

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|_2 \quad (17.8)$$

na množině \mathbb{R}^n . Tato funkce je konvexní a po zavedení pomocných proměnných z_i (podobná úprava jako v §12.1.1) lze její minimalizování formulovat jako SOCP:

$$\begin{aligned} \min \quad & z_1 + \dots + z_m \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2 \leq z_i, \quad i = 1, \dots, m. \end{aligned}$$

Pro případ $n = 2$ má úloha jednoduchý mechanický model¹. Do vodorovného prkna vyvrátáme 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 $f(\mathbf{x})$. ♦

17.4.5 Semidefinitní programování (SDP)

Věta 17.2. 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 $\mathbf{x}^T(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x} + \beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$. ■

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.

Konvexní kužel je konvexní množina. To umožňuje formulovat třídu konvexních úloh známou jako *semidefinitní programování* (SDP). Jednou z možných formulací je²

$$\begin{aligned} \min \quad & \langle \mathbf{C}, \mathbf{X} \rangle \\ \text{za podmíněk} \quad & \mathbf{X} \text{ je pozitivně semidefinitní} \\ & \langle \mathbf{A}_i, \mathbf{X} \rangle = b_i, \quad i = 1, \dots, m. \end{aligned} \quad (17.9)$$

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 $\langle \mathbf{A}, \mathbf{X} \rangle = \sum_i \sum_j a_{ij} x_{ij}$ označuje skalární součin matic (viz §2.1.7).

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 (17.9) se redukuje na LP. V tom případě v součinech $\langle \mathbf{C}, \mathbf{X} \rangle$ a $\langle \mathbf{A}_i, \mathbf{X} \rangle$ nediagonální prvky matice \mathbf{X} nehrají žádnou roli. Diagonální matice je pozitivně semidefinitní, právě když všechny její prvky jsou nezáporné (viz Cvičení 6.16). Tedy úloha (17.9) 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 17.14. *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 (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 §4.6.2). Místo součinu vzdáleností maximalizujeme jeho logaritmus (což je rostoucí funkce), tedy funkci

$$f(\mathbf{x}) = \log \prod_{i=1}^m (\mathbf{a}_i^T \mathbf{x} - b_i) = \sum_{i=1}^m \log(\mathbf{a}_i^T \mathbf{x} - b_i). \quad (17.10)$$

Tato funkce je konkávní. Funkce je definována jen pro vnitřní body mnohostěnu (tj. body splňující $\mathbf{A}\mathbf{x} > \mathbf{b}$) 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. ♦

17.5 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

$$Y \supseteq X \implies \min_{\mathbf{x} \in Y} f(\mathbf{x}) \leq \min_{\mathbf{x} \in X} f(\mathbf{x}). \quad (17.11)$$

Jak toho použít? Předpokládejme, že účelová funkce f je jednoduchá (např. konvexní), ale množina X přípustných řešení je komplikovaná a úloha $\min_{\mathbf{x} \in X} f(\mathbf{x})$ je díky tomu obtížná.

²Namítnete, že nemůžeme mluvit o konvexitě úlohy (17.9), protože v této úloze optimalizujeme přes množinu matic a konvexitu jsme definovali pro množiny a funkce vektorů. Definice konvexitu 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 lineárním prostoru (viz učebnice lineární algebry). Podobně již pro Větu 17.2.

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 (17.11) 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 pak o **konvexní relaxaci**. Příkladem konvexní relaxace je LP relaxace, kterou jsme už viděli: lineární program (12.24) je konvexní relaxace lineárního celočíselného programu (12.23).

17.6 Cvičení

17.1. 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?

17.2. Dokažte, že množina optimálních řešení konvexní optimalizační úlohy je konvexní.

17.3. Významnou vlastností konvexních funkcí je to, že každé lokální minimum funkce je zároveň globální (Věta 17.1). 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?

17.4. Najděte spojitou funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$, jejíž každé lokální minimum je zároveň globální a

- a) není konvexní
- b) (*) žádná její subkontura není konvexní množina.

17.5. Chceme rozestavit n lidí v místnosti čtvercového půdorysu tak, aby ‘každý byl od každého co nejdále’. Navrhněte možné formulace této úlohy a u každé určete, zda je konvexní.

17.6. (*) 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)

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{y} + dz \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{y} \geq \mathbf{b}z \\ & \mathbf{e}^T \mathbf{y} + fz = 1 \\ & z \geq 0 \end{aligned}$$

17.7. 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) $\min\{\mathbf{x}^T \mathbf{B} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1\}$

17.8. (*) Formulujte úlohu $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A} \mathbf{x} - \mathbf{b}\|_4$ jako konvexní QCQP.

17.9. (*) 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.

17.10. Sedumi (<http://sedumi.ie.lehigh.edu>) je jednou z implementací algoritmu na řešení úloh LP, QP, SOCP a SDP (používá tzv. algoritmus vnitřního bodu). Je k dispozici i pro Matlab. Knihovnu si stáhněte a pak prostudujte a pochopte nápovědu k funkci `sedumi.m`, která je matlabským rozhraním knihovny.

Nápověda a řešení

17.3. 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.

17.5. Necht' $\mathbf{x}_i \in \mathbb{R}^2$ jsou souřadnice i tého člověka v místnosti. Bez ztráty obecnosti necht' je místnost množina $[-1, 1] \times [-1, 1]$ (kde $[-1, 1]$ značí uzavřený interval). Jedna přirozená formulace je taková, že hledáme takové $\mathbf{x}_1, \dots, \mathbf{x}_n$ ležící v tomto čtverci, aby číslo $\min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|$ bylo maximální. Tato úloha není konvexní, neboť její účelová funkce není konvexní.

17.6.a) Ne.

17.6.b) Uvažujte substituci $\mathbf{y} = \mathbf{x}/(\mathbf{e}^T \mathbf{x} + f)$, $z = 1/(\mathbf{e}^T \mathbf{x} + f)$.

17.7.a) Viz Cvičení 12.6.

17.7.b) Substituuje $\mathbf{y} = \mathbf{x} - \mathbf{b}$.

17.7.c) Optimální hodnota je nula.

17.9. 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 18

(★) Lagrangeova dualita

Zatímco v §15 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.

18.1 Minimální nerovnost

Pro libovolné množiny X a Y a libovolnou funkci $L: X \times Y \rightarrow \mathbb{R}$ platí **minimaxní nerovnost**

$$\min_{x \in X} \underbrace{\max_{y \in Y} L(x, y)}_{F(x)} \geq \max_{y \in Y} \underbrace{\min_{x \in X} L(x, y)}_{G(y)}. \quad (18.1)$$

Zde předpokládáme, že všechna minima a maxima existují¹. V nerovnosti (18.1) nastává rovnost, právě když existuje bod $(x^*, y^*) \in X \times Y$ takový, že

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall x \in X, y \in Y. \quad (18.2)$$

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 (18.2) lze psát jako $F(x^*) = G(y^*)$.

Příklad 18.1. Necht' $X = Y = \{1, 2, 3, 4\}$.

	1	2	3	4	$F(x)$
1	-1	4	7	4	7
2	4	4	6	-2	6
3	1	5	3	3	5
4	3	5	3	2	5
$G(y)$	-1	4	3	-2	

	1	2	3	4	$F(x)$
1	-1	4	7	4	7
2	4	4	6	-2	6
3	1	0	3	3	3
4	3	3	3	2	3
$G(y)$	-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 $\min_{x \in X} F(x) = \max_{y \in Y} G(y)$. ♦

¹Kdyby ne, mohli bychom min/max nahradit inf/sup a nerovnost by stále platila.

18.2 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 (18.1) 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

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}). \quad (18.3)$$

Podívejme se, jak po této volbě vypadá levá strana nerovnosti (18.1), tedy primární úloha. Je

$$F(\mathbf{x}) = \max_{\mathbf{y} \in Y} L(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{y} \geq \mathbf{0}} [f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x})] = \begin{cases} f(\mathbf{x}) & \text{když } \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ \infty & \text{jinak,} \end{cases} \quad (18.4)$$

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 (18.4) do levé strany (18.1) tedy

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} L(\mathbf{x}, \mathbf{y}) = \min \{ f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \}. \quad (18.5)$$

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

$$\max_{\mathbf{y} \in Y} \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{y} \geq \mathbf{0}} G(\mathbf{y}), \quad (18.6)$$

kde

$$G(\mathbf{y}) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x} \in X} [f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x})]. \quad (18.7)$$

Přesnější tvar duální úlohy nezískáme, dokud nebudeme mít zadánu primární úlohu (tedy X, f, \mathbf{g}) konkrétněji. Nyní můžeme říci jen to, že úloha (18.7) 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 16.8 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 (18.6) se nazývá **Lagrangeova duální úloha** k (primární) úloze (18.5). Funkce (18.3) se říká Lagrangeova funkce – všimněte si, že je to ta samá funkce, kterou jste potkali u metody Lagrangeových multiplikátorů.

18.3 Silná dualita

Z nerovnosti (18.1) plyne, že optimální hodnota primární úlohy není menší než optimální hodnota duální úlohy. Tato skutečnost je známa jako věta o **slabé dualitě**. Rozdílu mezi primární a duální optimální hodnotou se říká **dualitní mezera**. Když v nerovnosti (18.1) 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 18.1, 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

$$g_1(\mathbf{x}) < 0, \dots, g_m(\mathbf{x}) < 0. \quad (18.8)$$

Pokud je prvních $k \leq m$ funkcí g_i afinních, podmínku (18.8) lze změkčit na

$$g_1(\mathbf{x}) \leq 0, \dots, g_k(\mathbf{x}) \leq 0, g_{k+1}(\mathbf{x}) < 0, \dots, g_m(\mathbf{x}) < 0. \quad (18.9)$$

Věta 18.1. *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 (18.5) a (18.6) jsou si rovny.

Dále uvedme obdobu věty o komplementaritě.

Věta 18.2. *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*

$$y_i g_i(x_i) = 0 \quad \forall i = 1, \dots, m. \quad (18.10)$$

Důkaz. Platí

$$F(\mathbf{x}) = f(\mathbf{x}) = G(\mathbf{y}) = \min_{\mathbf{x}' \in X} [f(\mathbf{x}') + \mathbf{y}\mathbf{g}(\mathbf{x}')] \leq f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}) \leq f(\mathbf{x}).$$

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

$$\mathbf{y}^T \mathbf{g}(\mathbf{x}) = 0. \quad (18.11)$$

To je ale ekvivalentní podmínkám (18.10), protože $\mathbf{y} \geq \mathbf{0}$ a $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. ■

18.4 Příklady

Příklad 18.2. 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

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{\mathbf{y} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{y}) = \min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b} \}. \quad (18.12)$$

Odvodíme duální úlohu. Lagrangeova funkce je

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{c}^T \mathbf{x} + \mathbf{y}^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = (\mathbf{c}^T - \mathbf{y}^T \mathbf{A})\mathbf{x} + \mathbf{y}^T \mathbf{b}, \quad (18.13)$$

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

$$\max_{\mathbf{y} \geq \mathbf{0}} G(\mathbf{y}) = \max\{ \mathbf{b}^T \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^m, \mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0} \}. \quad (18.14)$$

To je stejný výsledek, jaký bychom dostali podle předpisu (15.1) na konstrukci duálního LP. Konkrétně, primární a duální úloha je dvojice úloh (15.2). \blacklozen

Příklad 18.3. 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

$$\min_{\mathbf{x} \in \mathbb{R}_+^n} \max_{\mathbf{y} \geq \mathbf{0}} L(\mathbf{x}, \mathbf{y}) = \min\{ \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}. \quad (18.15)$$

Lagrangeova funkce je (18.13). 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

$$\max_{\mathbf{y} \geq \mathbf{0}} G(\mathbf{y}) = \max\{ \mathbf{b}^T \mathbf{y} \mid \mathbf{A}^T \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0} \}. \quad (18.16)$$

To je opět stejný výsledek, jaký bychom dostali podle předpisu (15.1). \blacklozen

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

$$L(\mathbf{x}, \mathbf{y}_+, \mathbf{y}_-) = f(\mathbf{x}) + \mathbf{y}_+^T \mathbf{g}(\mathbf{x}) - \mathbf{y}_-^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + (\mathbf{y}_+ - \mathbf{y}_-)^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \mathbf{y}^T \mathbf{g}(\mathbf{x}), \quad (18.17)$$

kde jsme označili $\mathbf{y} = \mathbf{y}_+ - \mathbf{y}_-$. Funkci (18.17) můžeme nyní přejmenovat na $L(\mathbf{x}, \mathbf{y})$, která má stejný tvar jako (18.3). 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

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in \mathbb{R}^m} L(\mathbf{x}, \mathbf{y}) = \min\{ f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) = \mathbf{0} \}, \quad (18.18)$$

jak jsme chtěli. Duální úloha je

$$\max_{\mathbf{y} \in \mathbb{R}^m} \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{y} \in \mathbb{R}^m} G(\mathbf{y}). \quad (18.19)$$

Všimněte si, že Slaterova podmínka nebude platit, když zobrazení \mathbf{g} nebude afinní.

Příklad 18.4. Napišme duální úlohu k úloze

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

Máme $X = \mathbb{R}^n$ a

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{x} + \mathbf{y}^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{A}\mathbf{x} + \mathbf{y}^T \mathbf{b},$$

duální funkce je tedy

$$G(\mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{x}^T \mathbf{x} - \mathbf{y}^T \mathbf{A}\mathbf{x} + \mathbf{y}^T \mathbf{b}).$$

Řešení musí splňovat $\partial L(\mathbf{x}, \mathbf{y})/\partial \mathbf{x} = \mathbf{0}$, což dá $\mathbf{x} = \mathbf{A}^T \mathbf{y}/2$. Po dosazení dostaneme

$$G(\mathbf{y}) = L(\mathbf{A}^T \mathbf{y}/2, \mathbf{y}) = \mathbf{b}^T \mathbf{y} - \mathbf{y}^T \mathbf{A} \mathbf{A}^T \mathbf{y}/4. \quad \blacklozenge$$

Příklad 18.5. Řešme lineární program

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}.$$

Duál bychom mohli sestavit podle návodu v kapitole o dualitě v LP. Ale postupujme jinak. Zvolme Lagrangeovu funkci jako (18.12) 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). \blacklozenge

Příklad 18.6. Řešme celočíselný lineární program

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}.$$

Necht' Lagrangeova funkce je (18.12) a necht' $X = \{0, 1\}^n$. Duální úloha je stejná jako v minulém příkladě. Zde předpoklady Věty 18.1 neplatí, protože množina X není konvexní. Opravdu, silná dualita u celočíselného programování obecně neplatí. \blacklozenge

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 18.7. Úloha QCQP

$$\min\{\mathbf{x}^T \mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{x} = 1\} \quad (18.20)$$

není konvexní, pokud matice \mathbf{A} není pozitivně semidefinitní. Řešení ale najdeme snadno pomocí spektrálního rozkladu:

$$\min_{\mathbf{x}^T \mathbf{x} = 1} \mathbf{x}^T \mathbf{A}\mathbf{x} = \min_{\mathbf{x}^T \mathbf{x} = 1} \mathbf{x}^T \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \mathbf{x} = \min_{\mathbf{z}^T \mathbf{z} = 1} \mathbf{z}^T \mathbf{\Lambda} \mathbf{z} = \lambda_{\min}(\mathbf{A}).$$

Ukážeme navíc, že platí silná dualita. Máme

$$L(\mathbf{x}, y) = \mathbf{x}^T \mathbf{A}\mathbf{x} + y(1 - \mathbf{x}^T \mathbf{x}) = \mathbf{x}^T (\mathbf{A} - y\mathbf{I})\mathbf{x} + y.$$

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 6.5) $\lambda_{\min}(\mathbf{A}) \geq y$. Duální úloha tedy je

$$\max\{y \in \mathbb{R} \mid \lambda_{\min}(\mathbf{A}) \geq y\}.$$

Ta má zřejmé optimální řešení $y = \lambda_{\min}(\mathbf{A})$. \blacklozenge

Kapitola 19

Vícekriteriální optimalizace

19.1 Uspořádání na množině

Binární relace na množině Y je množina $R \subseteq Y \times Y$. Binární relace je

- **reflexivní**, když $(x, x) \in R$ pro každé $x \in Y$,
- **transitivní**, když $(x, y) \in R$ a $(y, z) \in R$ implikuje $(x, z) \in R$,
- **antisymetrická**, když $(x, y) \in R$ a $(y, x) \in R$ implikuje $x = y$.

Částečné uspořádání (krátce jen **uspořádání**) na množině Y je binární relace na Y , která je reflexivní, transitivní a antisymetrická.

Relace uspořádání se obvykle značí infixově symbolem \preceq , tedy místo $(x, y) \in R$ píšeme $x \preceq y$. Pokud potřebujeme rozlišit více různých kvasi-uspořádání, používáme symboly jako $\leq_1, \leq_2, \leq', \preceq'$ atd.

Prvky $x, y \in Y$ jsou **srovnatelné** v uspořádání \preceq , když $x \preceq y$ nebo $y \preceq x$ nebo obojí. (Kvasi-)uspořádání je **úplné** (neboli **totální**), když každé dva prvky z Y jsou srovnatelné.

Příklad 19.1.

- $Y = \mathbb{R}$ a \preceq je přirozené uspořádání reálných čísel. Tato relace je úplné uspořádání.
- $Y \subseteq 2^U$ a \preceq je relace inkluze na množině 2^U , tedy $x \preceq y$ právě když $x \subseteq y$. Zde U je libovolná množina a 2^U značí množinu všech jejích podmnožin. Tato relace je uspořádání, ale není úplné.
- $Y = \mathbb{N}$ a \preceq je relace dělitelnosti, tj. $x \preceq y$ právě když x dělí y . Tato relace je uspořádání, není úplné.
- $\mathbf{x} \preceq \mathbf{y}$ právě když $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$. Tato relace není antisymetrická, tedy není uspořádání.



Nás ovšem nejvíce zajímá případ $Y = \mathbb{R}^n$.

Příklad 19.2. Příklady uspořádání na množině \mathbb{R}^n :

- Uspořádání 'po složkách': $\mathbf{x} \preceq \mathbf{y}$ právě když $x_i \leq y_i$ pro všechna $i = 1, \dots, n$. Toto uspořádání není úplné: např. pro $n = 2$ jsou vektory $\mathbf{x} = (0, 1)$ a $\mathbf{y} = (1, 0)$ nesrovnatelné.
- Lexikografické ('slovníkové') uspořádání: $\mathbf{x} \preceq \mathbf{y}$ právě když buď $\mathbf{x} = \mathbf{y}$, nebo existuje m takové, že $x_m < y_m$ a pro všechna $i < m$ platí $x_i = y_i$.

Toto uspořádání je úplné.

- Definujme $\mathbf{x} \preceq \mathbf{y}$ právě když $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$. Tato relace není uspořádání, protože není antisymetrická. \blacklozenge

Prvek $x \in Y$ se nazývá (vzhledem ke kvasi-uspořádání \preceq)

- **minimální prvek** množiny Y , když $y \preceq x$ implikuje $x \preceq y$, pro všechna $y \in Y$.
- **nejmenší prvek** množiny Y , když $x \preceq y$, pro všechna $y \in Y$.

Pro totální kvasi-uspořádání oba pojmy splývají.

19.2 Úlohy vícekritériální optimalizace

V klasické optimalizaci jsme se zabývali úlohami typu

$$\min_{x \in X} f(x), \quad (19.1)$$

kde X je množina přípustných řešení a $f: X \rightarrow \mathbb{R}$ je účelová (neboli kritériální) funkce. Optimální hodnoty této úlohy jsou minimální prvky množiny $f(X) = \{f(x) \mid x \in X\} \subseteq \mathbb{R}$. Zde pojem 'minimální prvek' se myslí vzhledem k přirozenému uspořádání na \mathbb{R} .

Zobecněme tuto úlohu. Necht' $f: X \rightarrow Y$ a necht' \preceq je (kvasi-)uspořádání na množině Y . Pak úlohou (19.1) budeme rozumět nalezení minimálních prvků množiny $f(X) \subseteq Y$ vzhledem k uspořádání \preceq . Případně pro každý minimální prvek y množiny $f(X)$ chceme najít argument $x \in X$, ve kterém se nabývá, tedy splňujícím $y = f(x)$.

Nejčastěji v aplikacích potkáme případ $Y = \mathbb{R}^n$. Pak mluvíme o *vícekritériální optimalizaci*¹, neboť vlastně chceme minimalizovat více skalárních kritérií (složek zobrazení $\mathbf{f}: X \rightarrow \mathbb{R}^n$, hodnoty zobrazení jsou vektory a tedy ho píšeme tučně) najednou. Dále se omezíme pouze na tento případ.

Příklad 19.3. V obchodě nabízejí čtyři druhy aut s těmito vlastnostmi:

		VW Golf	Opel Astra	Ford Focus	Toyota Corolla
cena	[tis. euro]	16	15	14	15
spotřeba	[l/100km]	7.2	7.0	7.5	8.2

Chceme levné auto s malou spotřebou. Která auta je dobré si koupit a která naopak nekoupit?

Máme $X = \{\text{VW Golf, Opel Astra, Ford Focus, Toyota Corolla}\}$ a $Y = \mathbb{R}^2$. Tabulka definuje zobrazení \mathbf{f} . Není ovšem jasné, jaké (kvasi-)uspořádání na množině \mathbb{R}^2 použít pro rozhodování.

Rozhodujme se vzhledem k uspořádání 'po složkách' \leq^2 . Vzhledem k tomuto uspořádání nemá množina $\mathbf{f}(X)$ nejmenší prvek, neboli kritéria jsou v konfliktu. Její minimální prvky jsou $\mathbf{f}(\text{Opel Astra}) = (15, 7.0)$ a $\mathbf{f}(\text{Ford Focus}) = (14, 7.5)$.

Rozhodujme se vzhledem k lexikografickému uspořádání, přesněji nejprve se rozhodujme dle ceny a pak dle spotřeby. Nyní minimální prvek množiny $\mathbf{f}(X)$ je $\mathbf{f}(\text{Ford Focus}) = (14, 7.5)$. \blacklozenge

Příklad 19.4. Chceme řešit (přesně či přibližně) nehomogenní lineární soustavu $\mathbf{Ax} \approx \mathbf{b}$, ale zároveň chceme, aby velikost vektoru \mathbf{x} byla malá. Řešíme tedy úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{Ax} - \mathbf{b}\|_2, \|\mathbf{x}\|_2).$$

¹Angl. *multicriteria optimization*. Názvosloví ovšem není jednotné, jindy se používají názvy *multiobjective optimization* nebo *vector optimization* (neboť hodnoty zobrazení f jsou vektory).

Jaké jsou minimální prvky množiny $\mathbf{f}(\mathbb{R}^2)$ vzhledem k uspořádání po složkách? Je jich nekonečně mnoho, obrázek. Vezmeme-li lexikografické uspořádání a je-li soustava přeuročena, dostaneme metodu nejmenší normy. ♦

Příklad 19.5. V okrese je n vesnic se souřadnicemi $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^2$. Do jakého místa $\mathbf{x} \in \mathbb{R}^2$ máme umístit heliport, aby byl ke všem vesnicím blízko?

Máme $X = \mathbb{R}^2$, $Y = \mathbb{R}^n$, a $f_i(\mathbf{x}) = \|\mathbf{a}_i - \mathbf{x}\|_2$. Řešíme úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^2} (\|\mathbf{a}_1 - \mathbf{x}\|_2, \dots, \|\mathbf{a}_n - \mathbf{x}\|_2).$$

Součástí úlohy není (kvasi-)uspořádání na množině \mathbb{R}^n . Jaká (kvasi-)uspořádání jsou vhodná?

Uspořádání po složkách: množina minimálních prvků množiny $\mathbf{f}(\mathbb{R}^2)$ je konvexní obal bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$. To je intuitivně jasné (i když přesný důkaz vynecháme): neleží-li \mathbf{x} v tomto konvexním obalu, můžeme změnit \mathbf{x} tak, že se vzdálenost k některým bodům \mathbf{a}_i zmenší a k ostatním se nezvětší. Tedy je hloupost umístit heliport mimo tento konvexní obal.

Max-uspořádání: vede na úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^2} \max_{i=1}^n \|\mathbf{a}_i - \mathbf{x}\|_2.$$

Toto je úloha klasické (skalární) optimalizace. Této formulaci se někdy říká *minimaxní* formulace. Minimalizujeme vzdálenost heliportu od nejvzdálenější vesnice. ♦

Rejstřík

afinní

- funkce, 48
- kombinace, 44, 182
- nezávislost, 47
- obal, 44, 182
- podprostor, 44, 182
- zobrazení, 47

antisymetrická matice, 19

argument minima, 4

báze

- lineárního podprostoru, 33
- mnohostěnu, 193
- sousední, 194
- standardní, 22

bázová proměnná, 193

bázové řešení, 193

bijektivní zobrazení, 3

binární relace, 2

bloková matice, 22

bod, 46

- extremální, 184
- hraniční, 127
- sedlový, 135, 238
- stacionární, 135
- vnitřní, 127

certifikát optimality, 208

cyklení, 199

definitnost matice na podprostoru, 151

derivace

- funkce jedné proměnné, 112
- jednostranná, 112
- oboustranná, 112
- parciální, 113
- směrová, 119
- totální, 114
- zprava/zleva, 112

determinant, 19

diferencovatelnost, 114

- funkce jedné proměnné, 112
- spojitá, 114
- totální, 114
- zobrazení
 - ve směru, 119

dimenze

- afinního podprostoru, 46
- lineárního podprostoru, 33
- mnohostěnu, 183
- stěny mnohostěnu, 186

doplnění na čtverec, 88

dualita

- silná, 207
- slabá, 206, 240
- v lineárním programování, 205

dyáda, 22

elipsoid, 89

epigraf, 221

extrém, 4

- globální, 128
- vázaný rovnostmi, 148
- vázaný, 6, 130
- volný, 6, 130

extremální bod, 184

faseta, 186

forma

- kvadratická, 83
- lineární, 36

Frobeniova

- norma, 21

funkce, 2

- afinní, 48
- konkávní, 219
- konvexní, 219
- kvadratická, 87
- Lagrangeova, 154
- lineární, 36

po částech afinní, 164
 účelová, 6

gradient, 120
 graf funkce, 110
 Gramm-Schmidtova ortonormalizace, 55

hodnost, 38
 plná, 38
 hrana mnohostěnu, 186
 hraniční bod množiny, 127
 hranice množiny, 127

idempotentní matice, 41
 idempotentní transformace, 3
 injektivní zobrazení, 3
 invertovatelná
 matice, 19
 inverze matice, 19
 inverzní zobrazení, 3
 involuce, 3
 isometrie, 54

jednotková matice, 17
 jednotková sféra, 169

kolmý, 52
 kombinace
 afinní, 44, 182
 konvexní, 181, 182
 lineární, 32, 182
 nezáporná, 182
 komplementární
 podprostory, 34
 projektory, 43
 komplementarita, 206
 konvexní
 funkce, 219
 kombinace, 182
 kužel, 182
 množina, 181, 182
 mnohostěn, 183
 obal, 181, 182
 optimalizační úloha, 229
 Kroneckerovo delta, 17
 kuželosečka, 90
 kvadratická
 forma, 83
 funkce, 87

kvadrika, 89

Lagrangeův multiplikátor, 154
 Lagrangeova funkce, 154
 Lagrangeovy multiplikátory, 150
 Laplacián grafu, 92
 lineární
 program v rovnicovém tvaru, 163
 forma, 36
 funkce, 36
 kombinace, 32, 182
 nerovnice, 162
 nezávislost, 32
 obal, 32, 182
 optimalizace, 162
 podprostor, 33, 182
 program, 162
 rovnice, 162
 soustava, 23, 66
 soustava nehomogenní nedourčená, 66
 soustava homogenní, 24
 soustava homogenní přeurená, 100
 soustava nehomogenní, 24, 66
 soustava nehomogenní přeurená, 66
 zobrazení, 35
 lineární varieta, 47
 LP relaxace, 172

matice, 17
 antisymetrická, 19
 bloková, 22
 čtvercová, 17
 diagonální, 17
 diagonalizovatelná, 80
 Hessova, 122
 idempotentní, 41
 indefinitní, 84
 invertovatelná, 19
 inverzní, 19
 Jacobiho, 114
 jednotková, 17
 nulová, 17
 obdélníková, 17
 ortogonální, 54
 permutační, 55
 pozitivně/negativně (semi)definitní, 84
 regulární, 19
 rotační, 54

- singulární, 19
- široká, 17
- speciální ortogonální, 54
- symetrická, 19
- transponovaná, 18
- trojúhelníková, 17
- úzká, 17
- metoda
 - Gauss-Newtonova, 142
 - Levenberg-Marquardtova, 144
 - čtverců, 66
 - Newtonova, 139
 - simplexová, 192
 - simplexová dvoufázová, 201
 - simplexová základní, 197
- metrika eukleidovská, 52
- minimální
 - hodnota funkce na množině, 4
 - prvek množiny, 4
- minimum
 - funkce na množině, 4, 128
 - funkce na množině ostré, 4
 - globální, 128
 - lokální, 130
 - lokální ostré, 130
 - ostré, 128
- množina
 - číselná, 2
 - konvexní, 181, 182
 - omezená, 127
 - otevřená, 127
 - uzavřená, 127
- mnohostěn konvexní, 183
- mocnina matice, 18
- monom, 78
- nadrovina, 46, 183
 - opěrná, 186
- nejmenší
 - čtverce, 66
 - norma, 71
 - prvek množiny, 4
- nerovnost
 - Jensenova, 219
 - minimaxní, 238
 - trojúhelníková, 52
- nezávislost
 - afinní, 47
 - lineární, 32
- norma
 - p -norma, 169
 - eukleidovská, 52
 - Frobeniova, 21
 - matice, 21
 - vektorová, 169
- nulita matice, 40
- nulová matice, 17
- obal
 - afinní, 44, 182
 - konvexní, 182
 - lineární, 32, 182
 - nezáporný, 182
- omezení, 6
- omezená množina, 127
- opěrná nadrovina, 186
- optimální
 - hodnota, 6
 - řešení, 6
- ortogonální
 - doplňěk, 57
 - matice, 54
 - podprostory, 57
 - projekce, 59
 - vektory, 52
- ortonormální množina vektorů, 53
- otevřená množina, 127
- přímka, 187
- permutační matice, 17
- pivot, 194
- pivotové pravidlo, 197
- počátek, 33
- podprostor, 33
 - afinní, 44, 182
 - lineární, 33, 182
 - triviální, 33
 - vlastní, 79
- podprostory
 - komplementární, 34
 - ortogonální, 57
- poloprostor, 183
- polynom, 78
 - charakteristický, 79
 - homogenní, 78

- Taylorův, 122
- přímka, 46
- projekce, 35, 41
 - šikmá, 35
- projekce ortogonální, 59
- projektor, 41
 - ortogonální, 59
- projektory
 - komplementární, 43
- prostor
 - lineární, 32
 - nulový, 37
 - obrazů, 37
 - vektorový, 32
- pseudoinverze
 - matice s l.n. řádky, 72
 - matice s l.n. sloupci, 68
 - obecné matice, 73
- regrese, 69
 - lineární, 69
 - robustní, 171
- regulární
 - matice, 19
- regularizace, 71
- relace
 - binární, 2
- řešení
 - bázové, 193
 - degenerované, 193
 - optimální, 6
 - přípustné, 6, 193
- restrikce zobrazení, 3
- řez zobrazení, 119
- rovina, 46
- rovnice
 - lineární, 162
 - normální, 67
- rozklad matice
 - podle hodnoty, 38
 - podle vlastních čísel, 80
 - QR, 56
 - spektrální, 80
- sestupné iterační metody, 136
- simplexová
 - metoda, 192
 - tabulka, 196
- singulární
 - číslo, 102
 - matice, 19
 - rozklad, 102
 - vektor, 102
- skalár, 18
- skalární součin, 22
 - matic, 20
- součin
 - vektorů, 52
- slacková proměnná, 164
- složené zobrazení, 3
- směr
 - Gauss-Newtonovův, 142
 - Newtonův, 141
- směr
 - sestupný, 134
 - vzestupný, 134
- součin
 - skalární s. matic, 20
- součin
 - maticový, 18
 - skalární, 22, 52
 - vnější, 22
- souřadnice, 33
- spektrální
 - rozklad, 80
- spektrum matice, 78
- stínové ceny, 209
- stacionární bod, 135
- stopa matice, 20
- subkontura, 221
- surjektivní zobrazení, 3
- SVD, 102
- Sylvestrovo kritérium, 85
- symetrická matice, 19
- tečný prostor k množině, 152
- transformace, 3
- transpozice, 18
- triviální podprostor, 33
- trojúhelníková matice, 17
- úloha
 - dopravní, 168
 - duální, 205
 - ekvivalentní, 231
 - konvexní, 229

konvexní ve standardním tvaru, 231
Lagrangeova duální, 239
lineárního programování, 162
na největší nezávislou množinu, 175
na nejmenší vrcholové pokrytí, 174
na optimální výrobní program, 167
primární, 205
přiřazovací, 173
shlukování, 12
směšovací, 167
spojité optimalizace, 5
uzavřená množina, 127

varieta
 algebraická, 47
 lineární, 47

vektor, 3, 21
 normalizovaný, 53
 řádkový, 21
 sloupcový, 21

vektory
 afinně nezávislé, 47
 lineárně nezávislé, 32
 ortogonální, 52
 ortonormální, 53

věta
 Frobeniova, 66

vlastní
 číslo, 78
 vektor, 78

vlastní podprostor, 79
vnitřek množiny, 127
vnitřní bod množiny, 127
vrchol, 186
vrstevnice funkce, 110
vychýlená hodnota, 171

základní podprostory matice, 38

zobrazení, 2
 afinní, 47
 lineární, 35
 spojité, 111