

Optimalizace

Elektronická skripta předmětu A4B33OPT.
Text je průběhu semestru doplňován a vylepšován.
Toto je verze ze dne **23. října 2014.**

Tomáš Werner



Katedra kybernetiky
Fakulta elektrotechnická
České vysoké učení technické

Obsah

1	Formulace optimalizačních úloh	6
1.1	Matematické značení	6
1.1.1	Množiny	6
1.1.2	Zobrazení	7
1.1.3	Funkce a zobrazení více reálných proměnných	7
1.1.4	Extrémy funkce na množině	8
1.2	Obecná úloha spojité optimalizace	9
1.3	Cvičení	10
2	Maticová algebra	12
2.1	Operace s maticemi	12
2.2	Transpozice a symetrie	13
2.3	Hodnost a inverze	13
2.4	Determinant	14
2.5	Matice s jedním sloupcem nebo jedním řádkem	14
2.6	Matice sestavené z bloků	15
2.7	Hrubé chyby při práci s maticemi	16
2.8	Cvičení	17
3	Linearita	20
3.1	Lineární podprostory	20
3.2	Lineární zobrazení	21
3.2.1	Prostor obrazů	22
3.2.2	Nulový prostor	23
3.2.3	Některé věty o prostoru obrazů a nulovém prostoru	23
3.3	Afinní podprostor a zobrazení	24
3.4	Cvičení	26
4	Ortogonalita	28
4.1	Standardní skalární součin	28
4.2	Ortogonální vektory	28
4.3	Ortogonální podprostory	29
4.4	Čtyři základní podprostory matice	30
4.5	Matice s ortonormálními sloupci	30
4.6	QR rozklad	31
4.6.1	(*) Gramm-Schmidtova ortonormalizace	32
4.7	Cvičení	33

5	Nehomogenní lineární soustavy	35
5.1	Přibližné řešení ve smyslu nejmenších čtverců	35
5.1.1	Řešení pomocí QR rozkladu	36
5.1.2	Více o ortogonální projekci	37
5.2	Řešení s nejmenší normou	38
5.2.1	Pseudoinverze obecné matice s plnou hodnotí	39
5.3	Použití	39
5.3.1	Lineární regrese	40
5.4	Cvičení	41
6	Spektrální rozklad a kvadratické funkce	44
6.1	Vlastní čísla a vektory	44
6.1.1	Spektrální rozklad	45
6.2	Kvadratická forma	46
6.3	Kvadratická funkce	47
6.4	Cvičení	49
7	Rozklad podle singulárních čísel (SVD)	52
7.1	Existence a výpočet SVD	53
7.2	Ortonormální báze základních podprostorů matice	54
7.3	Nejbližší matice nižší hodnoti	54
7.4	Prokládání bodů podprostorem	55
7.4.1	Zobecnění na afinní podprostor	56
7.5	Přibližné řešení homogenní soustavy	57
7.6	(★) Pseudoinverze obecné matice	57
7.7	Cvičení	59
8	Nelineární funkce a zobrazení	60
8.1	Spojitosť	61
8.2	Parciální derivace	61
8.3	Totální derivace	62
8.3.1	Derivace složeného zobrazení	63
8.3.2	Derivace maticových výrazů	65
8.4	Směrová derivace	66
8.5	Gradient	67
8.6	Parciální derivace druhého řádu	67
8.7	Taylorův polynom	68
8.8	Cvičení	70
9	Analytické podmínky na lokální extrémny	71
9.1	Vlastnosti podmnožin \mathbb{R}^n	71
9.2	Lokální extrémny	72
9.3	Volné lokální extrémny	72
9.4	Lokální extrémny vázané rovnostmi	74
9.4.1	Podmínky prvního řádu	74
9.4.2	(★) Podmínky druhého řádu	77
9.5	Lokální extrémny vázané nerovnostmi ‘hrubou silou’	79
9.6	Cvičení	80

10	Iterační algoritmy na volné lokální extrémy	84
10.1	Rychlost konvergence iteračních algoritmů	84
10.2	(★) Metoda zlatého řezu	85
10.3	Sestupné metody	86
10.4	Gradientní metoda	87
10.4.1	(★) Závislost na lineární transformaci souřadnic	87
10.5	Newtonova metoda	87
10.5.1	Použití na soustavy nelineárních rovnic	88
10.5.2	Použití na minimalizaci funkce	89
10.6	Gauss-Newtonova metoda	89
10.6.1	Rozdíl proti Newtonově metodě	91
10.6.2	Levenberg-Marquardtova metoda	91
10.7	Statistické odůvodnění kritéria nejmenších čtverců	92
10.8	Cvičení	92
11	Konvexita	93
11.1	Konvexní množiny	93
11.1.1	Čtyři kombinace a čtyři obaly	94
11.1.2	Operace zachovávající konvexitu množin	94
11.2	Konvexní funkce	95
11.2.1	Epigraf a subkontura	96
11.2.2	Konvexita diferencovatelných funkcí	98
11.2.3	Operace zachovávající konvexitu funkcí	99
11.3	Konvexní optimalizační úlohy	100
11.4	Cvičení	102
12	Lineární programování	104
12.1	Konvexní polyedry	104
12.1.1	Stěny konvexního polyedru	105
12.1.2	(★) Dvě reprezentace polyedru	106
12.2	Úloha lineárního programování	106
12.3	Různé tvary úloh LP	107
12.3.1	Po částech afinní funkce	109
12.4	Některé aplikace LP	110
12.4.1	Optimální výrobní program	110
12.4.2	Směšovací (dietní) problém	111
12.4.3	Dopravní problém	111
12.4.4	Distribuční problém	112
12.5	Přeurčené lineární soustavy	112
12.5.1	Vektorové normy	112
12.5.2	Přibližné řešení přeurčených lineárních soustav	113
12.5.3	Použití na robustní regresi	114
12.6	Cvičení	115

13 Simplexová metoda	118
13.1 Stavební kameny algoritmu	119
13.1.1 Přechod k sousední standardní bázi	120
13.1.2 Kdy je sousední bazové řešení přípustné?	120
13.1.3 Co když je celý sloupec nekladný?	121
13.1.4 Ekvivalentní úpravy účelového řádku	121
13.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?	122
13.2 Základní algoritmus	123
13.3 Inicializace algoritmu	125
13.3.1 Dvoufázová simplexová metoda	125
13.4 Cvičení	127
14 Dualita v lineárním programování	129
14.1 Konstrukce duální úlohy	129
14.2 Věty o dualitě	130
14.3 Stínové ceny	132
14.4 Příklady na konstrukci a interpretaci duálních úloh	133
14.5 Cvičení	136
15 Konvexní optimalizační úlohy	138
15.1 Třídy optimalizačních úloh	138
15.2 Příklady nekonvexních úloh	139
15.2.1 Celočíslné programování	140
15.2.2 Konvexní relaxace nekonvexních úloh	141
15.3 Cvičení	142
16 Vícekriteriální optimalizace	143
16.1 Uspořádání na množině	143
16.2 Úlohy vícekriteriální optimalizace	144

Kapitola 1

Formulace optimalizačních úloh

Osnova přednášky: Příklady optimalizačních úloh: prokládání bodů parabolou, dietní problém, visící řetěz, shlukování / facility location. Co mají úlohy společného; vysvětlit obecnou formulaci spojité opt, minimum fce na množině. Co mají úlohy rozdílného: omezení nebo ne; diferencovatelné nebo ne; optimalizuje ekonomika (dietní prob), příroda (řetěz), statistický princip (LS). Co brát nebudeme: kombinatorická opt, variační počet, multikriteriální opt.

1.1 Matematické značení

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

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

1.1.1 Množiny

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

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

Číselné množiny budeme značit takto:

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

1.1.2 Zobrazení

Zobrazení z množiny A do množiny B značíme

$$f: A \rightarrow B \tag{1.1}$$

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

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

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

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

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

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

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

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

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

1.1.3 Funkce a zobrazení více reálných proměnných

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

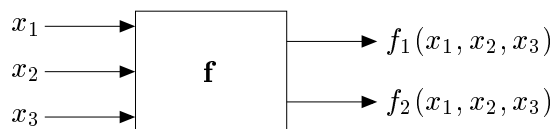
$$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m \tag{1.2}$$

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

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

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

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



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

1.1.4 Extrémy funkce na množině

Necht' $Y \subseteq \mathbb{R}$. Prvek $y \in Y$ nazveme *nejmenší prvek* (neboli *minimum*) množiny Y , jestliže $y \leq y'$ pro všechna $y' \in Y$. Nejmenší prvek značíme

$$\min Y.$$

Ne každá množina $Y \subseteq \mathbb{R}$ má nejmenší prvek (např. interval $(0, 1]$ ho nemá). Na druhou stranu je jasné, že Y nemůže mít více než jeden minimální prvek.

Mějme nyní funkci $f: X \rightarrow \mathbb{R}$, kde X je libovolná množina. Označme

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

obraz množiny X funkcí f . Pokud množina Y má nejmenší prvek, definujeme

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

a hovoříme o *minimu funkce f na množině X* . V tom případě existuje nejméně jeden prvek $x \in X$ tak, že $f(x) = \min Y$. Říkáme, že funkce *nabývá minimum* v bodě x . Pro podmnožinu množiny X , ve kterých se minimum nabývá, se používá symbol ‘argument minima’

$$\operatorname{argmin}_{x \in X} f(x) = \{ x \in X \mid f(x) = \min Y \}.$$

Podobně definujeme maximum funkce na množině. Minima a maxima funkce se souhrnně nazývají její *extrémy* nebo *optima*.

Jednodušeji můžeme říci, funkce $f: X \rightarrow \mathbb{R}$ nabývá svého minima na množině X v bodě $x \in X$, jestliže $f(x) \leq f(x')$ pro všechny $x' \in X$. Tato definice ovšem zamlčuje vztah k minimálnímu prvku množiny $Y \subseteq \mathbb{R}$.

Definuj zde take ostre extremy. Pozdeji definovat i ostre lokalni extremy.

Příklad 1.1.

- $\min_{x \in \mathbb{R}} |x - 1| = \min\{|x - 1| \mid x \in \mathbb{R}\} = \min \mathbb{R}_+ = 0, \operatorname{argmin}_{x \in \mathbb{R}} |x - 1| = \{1\}$

- Necht' $f(x) = \max\{|x|, 1\}$. Pak $\operatorname{argmin}_{x \in \mathbb{R}} f(x) = [-1, 1]$.

- Necht' $(a_1, a_2, \dots, a_5) = (1, 2, 3, 2, 3)$. Pak² $\max_{i=1}^5 a_i = 3, \operatorname{argmax}_{i=1}^5 a_i = \{3, 5\}$. □

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

1.2 Obecná úloha spojité optimalizace

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

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

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

$$g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \quad (1.3a)$$

$$h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l \quad (1.3b)$$

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

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

kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^l$ a $\mathbf{0}$ značí nulové vektory příslušné dimenze. Hledání minima funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině X se zapisuje také jako

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmíněk} \quad & g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \\ & h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l. \end{aligned} \quad (1.4)$$

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

Označme strany obdélníka x, y . Řešíme úlohu

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

neboli

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

Zde máme $n = 2, m = 2, l = 1$.

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

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

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

Příklad 1.3. Hledejme dvojici nejbližších bodů v rovině, z nichž jeden je na kružnici se středem v počátku a jednotkovým poloměrem a druhý je ve čtverci se středem v bodě $(2, 2)$ a jednotkovou stranou. Úlohu lze samozřejmě řešit snadno úvahou. Napišme ji ale ve tvaru (1.4).

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

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

Řešíme úlohu

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

V matematické analýze se řešením úlohy (1.4) říká *extrémy funkce f vázané podmínkami (1.3)*. Pokud omezení chybí, mluví se o *volných extrémech* funkce f . V matematické optimalizaci se vžilo poněkud odlišné názvosloví:

- Funkce f se nazývá *účelová* (také pokutová, cenová, kritériální) funkce.
- Prvky množiny X se nazývají *přípustná řešení*, což je vlastně protimluv, protože nemusí být řešeními úlohy (1.4). Prvkům množiny $\operatorname{argmin}_{\mathbf{x} \in X} f(\mathbf{x})$ se pak říká *optimální řešení*.
- Rovnice a nerovnice (1.3) se nazývají *omezující podmínky*, krátce *omezení*.
- Omezení (1.3a) příp. (1.3b) se nazývají omezení *typu nerovnosti* příp. *typu rovnosti*. Pokud omezení chybí ($m = l = 0$), jedná se o optimalizaci *bez omezení*.
- Pokud je množina X přípustných řešení prázdná (omezení si odporují), úloha se nazývá *nepřípustná*.
- Pokud se účelová funkce při splněných omezeních může zlepšovat nade všechny meze, úloha se nazývá *neomezená*.

1.3 Cvičení

Chtělo by to přidat cvičení na množinové značení, kartézský součin, zobrazení.

1.1. Vyřešte následující úlohy, přičemž slovní úlohy nejdříve formulujte ve tvaru (1.4). Stačí vám k tomu zdravý rozum a derivace funkcí jedné proměnné.

- $\min\{x^2 + y^2 \mid x > 0, y > 0, xy \geq 1\}$
- $\min\{(x - 2)^2 + (y - 1)^2 \mid x^2 \leq 1, y^2 \leq 1\}$
- Máte vyrobit papírovou krabici o objemu 72 litrů, jejíž délka je dvojnásobek její šířky. Jaké budou její rozměry, má-li se na ní spotřebovat co nejméně papíru? Tloušťka stěn je zanedbatelná.
- Jaké má rozměry válec s jednotkovým objemem a nejmenším povrchem?
- Najděte rozměry püllitru, na jehož výrobu je třeba co nejméně skla. Tloušťka stěn je zanedbatelná.

- f) Najděte obsah největšího obdélníka vepsaného do půlkruhu s poloměrem 1.
- g) Obdélník v rovině má jeden roh v počátku a druhý na křivce $y = x^2 + x^{-2}$, přičemž jeho strany jsou rovnoběžné se souřadnicovými osami. Pro jaké x bude jeho obsah minimální? Může být jeho obsah libovolně veliký?
- h) Najděte bod v rovině na parabole s rovnicí $y = x^2$ nejbližší bodu $(3, 0)$.
- i) Hektarová oblast obdélníkového tvaru se má obehnat ze tří stran živým plotem, který stojí 1000 korun na metr, a ze zbývajících strany obyčejným plotem, který stojí 500 korun na metr. Jaké budou rozměry oblasti při nejmenší ceně plotu?
- j) x, y jsou čísla v intervalu $[1, 5]$ takové, že jejich součet je 6. Najděte tato čísla tak, aby xy^2 bylo (a) co nejmenší a (b) co největší.
- k) Hledá se n -tice čísel $x_1, \dots, x_n \in \{-1, +1\}$ tak, že jejich součin je kladný a jejich součet minimální. Jako výsledek napište (co nejjednodušší) vzorec, který udává hodnotu tohoto minimálního součtu pro obecné n .
- l) *Potkaní biatlon*. Potkan stojí na břehu kruhového jezírka o poloměru 1 a potřebuje se dostat na protilehlý bod břehu. Potkan plave rychlostí v_1 a běží rychlostí v_2 . Chce se do cíle dostat co nejrychleji, přičemž může běžet, plavat, nebo zvolit kombinaci obojího. Jakou dráhu zvolí? Strategie potkana může být různá pro různé hodnoty v_1 a v_2 , vyřešte pro všechny kombinace těchto hodnot.

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

- a) $[-1, 0] \times \{1\}$
- b) $\mathbb{Z} \times \mathbb{Z}$
- c) $\mathbb{R} \times \mathbb{Z}$
- d) $(\mathbb{R} \times \mathbb{Z}) \cup (\mathbb{Z} \times \mathbb{R})$
- e) $\{(x, y) \mid x^2 + y^2 = 1\} \times \mathbb{R}$
- f) $\{(x, y) \mid x > 0, y > 0, xy = 1\}$
- g) $\{(x, y) \mid \min\{x, y\} = 1\}$
- h) $\{1/x \mid x \geq 1\}$
- i) $\{1/x \mid |x| \geq 1\}$
- j) $\{e^{-x^2} \mid x \in \mathbb{R}\}$
- k) $\{x + y \mid x^2 + y^2 < 1\}$
- l) $\{x - y \mid x^2 + y^2 = 1\}$
- m) $\{x + y \mid |x| \geq 1, |y| \geq 1\}$

Kapitola 2

Maticová algebra

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

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

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

Budeme používat tyto názvy:

- Pro $m = n$ se matice nazývá **čtvercová** a pro $m \neq n$ **obdélníková**, přičemž pro $m < n$ je **široká** a pro $m > n$ je **úzká**.
- **Diagonální prvky** matice jsou prvky a_{11}, \dots, a_{pp} , kde $p = \min\{m, n\}$. Matice je **diagonální**, když všechny nediagonální prvky jsou nulové (platí pro čtvercové i obdélníkové matice). Pokud \mathbf{A} je čtvercová diagonální ($m = n$), značíme $\mathbf{A} = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$.
- **Nulová matice** má všechny prvky nulové. Značíme ji $\mathbf{0}_{m \times n}$ (pokud jsou rozměry jasné z kontextu, pak pouze $\mathbf{0}$).
- **Jednotková matice** je čtvercová diagonální, jejíž diagonální prvky jsou jedničky. Značíme ji \mathbf{I}_n (pokud jsou rozměry jasné z kontextu, pak pouze \mathbf{I}).

2.1 Operace s maticemi

Na maticích jsou definovány následující operace:

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

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

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

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

Vlastnosti maticového součinu:

- $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ a $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- $\mathbf{AI}_n = \mathbf{A} = \mathbf{I}_m\mathbf{A}$
- $(\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B}) = \alpha(\mathbf{AB})$ (Mohli bychom si myslet, že výraz $\alpha\mathbf{A}$ je maticový součin, kde skalár $\alpha \in \mathbb{R}$ se považuje za matici rozměru 1×1 . Tak to ale není, protože vnitřní rozměr matic v součinu $\alpha\mathbf{A}$ by byl obecně různý.)

Obecně neplatí $\mathbf{AB} = \mathbf{BA}$ (čtvercové matice obecně nekomutují).

2.2 Transpozice a symetrie

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

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

Čtvercová matice se nazývá

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

2.3 Hodnost a inverze

Hodnost matice je velikost největší podmnožiny jejích lineárně nezávislých sloupců. Jinými slovy, je to dimenze lineárního obalu sloupců matice. Hodnost značíme $\text{rank } \mathbf{A}$. Platí (a není to snadné dokázat)

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

tedy místo pomocí sloupců lze hodnost ekvivalentně definovat pomocí řádků. Z toho plyne, pro každou matici je vždy

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

Pokud $\text{rank } \mathbf{A} = \min\{m, n\}$, říkáme, že matice má **plnou hodnost**. Čtvercová matice s plnou hodností se nazývá **regulární**, jinak je **singulární**. Uvedme ještě tvrzení o hodnosti součinu matic, které dokážeme později v §3.2.1:

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

Pokud matice $\mathbf{A} \in \mathbb{R}^{n \times m}$ a $\mathbf{B} \in \mathbb{R}^{m \times n}$ splňují

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

matice \mathbf{B} se nazývá **pravá inverze** matice \mathbf{A} a matice \mathbf{A} se nazývá **levá inverze** matice \mathbf{B} . Pravá či levá inverze nemusí existovat nebo nemusí být jediná.

Pravá inverze čtvercové matice existuje právě tehdy, když tato matice je regulární. Proto se regulární matici říká také **invertovatelná**. V tom případě je jediná a je rovna levé inverzi. Pak mluvíme pouze o **inverzi** matice a značíme ji \mathbf{A}^{-1} . Vlastnosti inverze:

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

2.4 Determinant

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

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

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

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

2.5 Matice s jedním sloupcem nebo jedním řádkem

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

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

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

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

Důležité jsou případy, kdy se v maticovém součinu vyskytují vektory:

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

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

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

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

Symbol $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ (jednička na i -tém místě) značí i -tý (sloupcový) vektor standardní báze. Dimenze n vektoru je určena kontextem. Příklad: pro $\mathbf{x} \in \mathbb{R}^n$ je $\mathbf{e}_i^T \mathbf{x} = x_i$.

2.6 Matice sestavené z bloků

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

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad [\mathbf{A} \quad \mathbf{B}], \quad \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}. \quad (2.7)$$

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

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

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

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

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

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

Příklad 2.2. Často je užitečné vnímat matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ jako matici sestavenou z bloků

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

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

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

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

Vyjádříme-li matici \mathbf{A} pomocí řádků a matici \mathbf{B} pomocí sloupců, je

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n] = [\mathbf{Ab}_1 \ \cdots \ \mathbf{Ab}_n] = \begin{bmatrix} \mathbf{a}_1^T \mathbf{B} \\ \vdots \\ \mathbf{a}_m^T \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{b}_1 & \cdots & \mathbf{a}_1^T \mathbf{b}_n \\ \vdots & \ddots & \vdots \\ \mathbf{a}_m^T \mathbf{b}_1 & \cdots & \mathbf{a}_m^T \mathbf{b}_n \end{bmatrix}. \quad (2.8)$$

Vyjádříme-li matici \mathbf{A} pomocí sloupců a matici \mathbf{B} pomocí řádků, je

$$\mathbf{AB} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_p] \begin{bmatrix} \mathbf{b}_1^T \\ \vdots \\ \mathbf{b}_p^T \end{bmatrix} = \mathbf{a}_1 \mathbf{b}_1^T + \cdots + \mathbf{a}_p \mathbf{b}_p^T. \quad (2.9) \quad \square$$

2.7 Hrubé chyby při práci s maticemi

Při manipulaci s maticovými výrazy a rovnicemi dělají studenti někdy hrubé chyby, kterých se lze při alespoň minimální pozornosti vyvarovat. Uvedme typické příklady těchto chyb.

Výraz je nesmyslný kvůli rozměrům matic

Jako první příklad uvedme chyby, kdy výraz nemá smysl kvůli rozměrům matic a vektorů. První typ těchto chyb porušuje syntaktická pravidla, např.:

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

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

- Zcela odstrašující příklad je použití ‘zlomku’ pro matice, např. $\frac{\mathbf{A}}{\mathbf{B}}$.

Ve druhém typu těchto chyb student napíše výraz nebo udělá závěr, který neodporuje syntaxi ale nedává smysl kvůli sémantice (významu), např.:

- Inverze čtvercové, ale evidentně singulární matice. Např. $(\mathbf{w}\mathbf{w}^T)^{-1}$, kde $\mathbf{w} \in \mathbb{R}^3$.
- Předpoklad existence levé inverze široké matice nebo pravé inverze úzké matice. Např. napíšeme $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$, kde $\mathbf{Q} \in \mathbb{R}^{5 \times 3}$.
- Tvrzení $\text{rank } \mathbf{A} = 5$, kde $\mathbf{A} \in \mathbb{R}^{3 \times 5}$. Je chybné, protože každá pětice vektorů z \mathbb{R}^3 je lineárně závislá.

Příklad 2.3. Vidíme-li výraz $(\mathbf{A}^T \mathbf{B})^{-1}$, musí nám okamžitě hlavou proběhnout tyto úvahy o rozměrech matic $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{k \times p}$:

- Abychom se vyhnuli syntaktické chybě v násobení, musí být $m = k$.
- Jelikož součin $\mathbf{A}^T \mathbf{B}$ je rozměru $n \times p$, musí být $n = p$, abychom se vyhnuli syntaktické chybě při inverzi. Teď tedy víme, že obě matice musí mít stejný rozměr.
- Z nerovnosti (2.4) je jasné, že pokud by \mathbf{A}^T byla úzká nebo \mathbf{B} široká, $\mathbf{A}^T \mathbf{B}$ by byla singulární a dostali bychom sémantickou chybu. Abychom se jí vyhnuli, musí být obě matice buď čtvercové nebo úzké, $m \geq n$.

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

Použití neexistujících maticových identit

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

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

Neekvivalentní úpravy (ne)rovníc

Zde student udělá chybný úsudek při *neekvivalentní úpravě* rovnice či nerovnice. Ekvivalentní a neekvivalentní úpravy skalárních rovnic známe již ze základní školy. Např. operace 'odmocni rovnici' je neekvivalentní, neboť sice $a = b$ implikuje $a^2 = b^2$, ale $a^2 = b^2$ neimplikuje $a = b$. Příklady:

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

Další nápady pro práci s maticemi

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

2.8 Cvičení

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

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

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

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

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

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

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

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

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

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 $\text{vec} \mathbf{A}$ je matice \mathbf{A} přerovnaná po sloupcích do vektoru. *Kroneckerův součin matic* (v Matlabu $\text{kron}(\mathbf{A}, \mathbf{B})$) je definován jako

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

Pro libovolné matice (s kompatibilními velikostmi) platí

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec} \mathbf{B}. \quad (2.10)$$

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

a) $\{\mathbf{b}_i^T \mathbf{Xa}_i = 0, i = 1, \dots, k\}$

b) $\mathbf{AX} + \mathbf{XA}^T = \mathbf{C}$

2.7. Součet diagonálních prvků čtvercové matice se nazývá její *stopa*.

a) Dokažte, že matice \mathbf{AB} a \mathbf{BA} mají stejnou stopu.

b) Dokažte, že rovnice $\mathbf{AB} - \mathbf{BA} = \mathbf{I}$ nemá řešení pro žádné \mathbf{A}, \mathbf{B} .

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

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

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

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

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

a) $\mathbf{A} + \mathbf{A}^T$ je symetrická

b) $\mathbf{A} - \mathbf{A}^T$ je antisymetrická

c) existuje symetrická \mathbf{B} a antisymetrická \mathbf{C} tak, že $\mathbf{A} = \mathbf{B} + \mathbf{C}$, přičemž \mathbf{B}, \mathbf{C} jsou dány jednoznačně.

d) $\mathbf{A}^T\mathbf{A}$ je symetrická

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

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

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

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

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

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

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

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

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

Nápověda ke cvičením

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

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

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

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

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

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

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

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

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

2.5.d) Rovnic je k^2 , neznámých je mn , kde $\mathbf{X} \in \mathbb{R}^{m \times n}$ a $\mathbf{A} \in \mathbb{R}^{k \times m}$. Je lineární.

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

Kapitola 3

Linearita

Množina $\mathbb{R}^{m \times n}$ matic pevného rozměru $m \times n$ spolu s operacemi $+$ (sčítání matic) a \cdot (násobení matice skalárem) tvoří lineární prostor nad tělesem reálných čísel. Speciálním případem je lineární prostor $\mathbb{R}^{n \times 1}$ jednosloupcových matic neboli, po ztotožnění popsaném v §2.5, lineární prostor \mathbb{R}^n všech n -tic reálných čísel. Zopakujte si z lineární algebry pojem lineárního prostoru!

3.1 Lineární podprostory

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

$$\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \tag{3.1}$$

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

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

V opačném případě jsou **lineárně závislé**. **Lineární obal** vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina

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

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

Množina $X \subseteq \mathbb{R}^n$ se nazývá **lineární podprostor** (častěji jen **podprostor**) lineárního prostoru \mathbb{R}^n , pokud libovolná lineární kombinace libovolných vektorů z X leží v X (říkáme, že množina X je uzavřená vůči lineárním kombinacím). **Báze** lineárního podprostoru $X \subseteq \mathbb{R}^n$ je lineárně nezávislá množina vektorů, jejíž lineární obal je X . Netriviální podprostor \mathbb{R}^n má nekonečný počet bází, každá báze má však stejný počet vektorů. Tento počet je **dimenzí** lineárního podprostoru, kterou značíme $\dim X$.

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

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

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

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

3.2 Lineární zobrazení

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

$$\mathbf{f}(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k) = \alpha_1 \mathbf{f}(\mathbf{x}_1) + \dots + \alpha_k \mathbf{f}(\mathbf{x}_k), \quad (3.3)$$

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

Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je lineární právě tehdy, existuje-li matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ taková, že

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}. \quad (3.4)$$

Důkaz jedné implikace je očividný: zobrazení (3.4) je lineární, neboť

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

což plyne z vlastností maticového součinu (viz §2.1). Důkaz obrácené implikace vynecháme, ale není obtížný. Lze navíc ukázat, že matice \mathbf{A} je zobrazením \mathbf{f} určena jednoznačně. Říkáme proto, že matice \mathbf{A} *reprezentuje* lineární zobrazení \mathbf{f} .

Příklad 3.5. Zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x_1, x_2) = (x_1 + x_2, x_1 - x_2, 2x_1)$ je lineární. To bychom dokázali ověřením podmínky (3.3). Ovšem je to patrné na první pohled, protože jej lze vyjádřit ve tvaru (3.4):

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

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

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = a_1 x_1 + \dots + a_n x_n, \quad (3.5)$$

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

Podívejme se blíže na vzorec (3.4). Výraz $\mathbf{A}\mathbf{x}$ je maticový součin matice $m \times n$ maticí $n \times 1$ (viz §2.5). Označíme-li $\mathbf{y} = \mathbf{A}\mathbf{x}$, je tedy podle (2.1)

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad (3.6)$$

neboli

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

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

$$\mathbf{Ax} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n, \quad (3.7)$$

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

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{x} \\ \vdots \\ \mathbf{a}_m^T \mathbf{x} \end{bmatrix}, \quad (3.8)$$

tedy složky vektoru \mathbf{Ax} jsou *skalární součiny řádků* matice a vektoru \mathbf{x} . Všimněte si, že (3.7) a (3.8) jsou speciální případy (2.8) a (2.9) z Příkladu 2.2.

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

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

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

Zde příklady lineárních zobrazení: rotace, zrcadlení, projekce, permutace, vektorový součin (oddělat příslušné cvičení); praktické příklady z Boyda.

3.2.1 Prostor obrazů

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

Prostor obrazů matice \mathbf{A} (neboli prostor obrazů zobrazení \mathbf{f}) je množina

$$\text{rng } \mathbf{A} = \{ \mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n \} \subseteq \mathbb{R}^m. \quad (3.9)$$

Interpretace prostoru obrazů:

- Je to množina $\mathbf{f}(\mathbb{R}^n)$ všech hodnot, jichž může zobrazení \mathbf{f} nabýt.
- Je to množina všech vektorů \mathbf{y} , pro které má lineární soustava $\mathbf{Ax} = \mathbf{y}$ řešení.
- Podle (3.7) je to lineární obal sloupců matice \mathbf{A} . Tedy je to lineární podprostor \mathbb{R}^m .

Z definice hodnosti je jasné, že

$$\dim \text{rng } \mathbf{A} = \text{rank } \mathbf{A}. \quad (3.10)$$

Věta 3.1. *Následující tvrzení jsou ekvivalentní:*

1. $\text{rng } \mathbf{A} = \mathbb{R}^m$
2. $\text{rank } \mathbf{A} = m$
3. Řádky matice \mathbf{A} jsou lineárně nezávislé.
4. Soustava $\mathbf{Ax} = \mathbf{y}$ má řešení pro každé \mathbf{y} .

5. Zobrazení f je surjektivní (viz §1.1.2).
6. Matice A má pravou inverzi, tj. existuje B tak, že $AB = I$.

Důkaz. Ekvivalence $1 \Leftrightarrow 2 \Leftrightarrow 3 \Leftrightarrow 4 \Leftrightarrow 5$ snadno plynou z definice hodnosti a z (3.10). Implikace $4 \Rightarrow 6$ platí, neboť soustava $Ab_i = e_i$ má řešení b_i pro každé i (kde e_i resp. b_i je i -tý sloupec matice I resp. B). Pro důkaz $6 \Rightarrow 4$ položíme $x = By$. \square

3.2.2 Nulový prostor

Nulový prostor matice A (také se nazývá *jádro* zobrazení f) je množina

$$\text{null } A = \{x \in \mathbb{R}^n \mid Ax = 0\} \subseteq \mathbb{R}^n. \quad (3.11)$$

Interpretace nulového prostoru:

- Je to množina všech vektorů, které se zobrazí na nulový vektor.
- Podle (3.8) je to množina všech vektorů, které jsou kolmé na každý řádek matice A . Z toho je vidět, že je to lineární podprostor \mathbb{R}^n .

Věta 3.2. *Následující tvrzení jsou ekvivalentní:*

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

Důkaz. Ekvivalence $1 \Leftrightarrow 3$ plyne z definice lineární nezávislosti (3.2), tj. $Ax = 0 \Rightarrow x = 0$. Ekvivalence $2 \Leftrightarrow 3$ plyne z definice hodnosti. Tvrzení 4 říká, že pro každé x, y je $Ax = Ay \Rightarrow x = y$, neboli $A(x - y) = 0 \Rightarrow x - y = 0$. To dokazuje $1 \Leftrightarrow 4$. Tvrzení 5 je ekvivalentní tomu, že matice A^T má pravou inverzi, tj. $A^T B^T = I$. Tedy $3 \Leftrightarrow 5$ plyne z Věty 3.1. \square

3.2.3 Některé věty o prostoru obrazů a nulovém prostoru

Následující věta je základní, její důkaz lze najít v každé učebnici lineární algebry.

Věta 3.3. *Pro každou matici $A \in \mathbb{R}^{m \times n}$ platí*

$$\dim \text{rng } A + \dim \text{null } A = n. \quad (3.12)$$

Neformální interpretace: Číslo $\dim \text{rng } A = \text{rank } A$ je dimenze podprostoru všech hodnot, které se mohou objevit na výstupu zobrazení. Číslo $\dim \text{null } A$ je dimenze podprostoru vstupních hodnot, které zobrazení 'smáčkne' do nulového vektoru. Věta 3.3 říká, že každá dimenze na vstupu se buď 'smáčkne' do nulového vektoru nebo se objeví na výstupu.

Dokažme nyní snadná tvrzení o prostoru obrazů a nulovém prostoru součinu matic.

Věta 3.4. *Pro libovolné matice A, B platí*

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

$$3. \text{null}(\mathbf{BA}) \supseteq \text{null } \mathbf{A}$$

$$4. \text{null}(\mathbf{BA}) = \text{null } \mathbf{A}, \text{ jestliže sloupce } \mathbf{B} \text{ jsou lineárně nezávislé}$$

Důkaz. Inkluze 1 říká, že jestliže soustava $\mathbf{z} = \mathbf{ABx}$ má řešení, pak i soustava $\mathbf{z} = \mathbf{Ay}$ má řešení. To ale platí, protože můžeme položit $\mathbf{y} = \mathbf{Bx}$. Rovnost 2 pak plyne z implikace $3 \Rightarrow 4$ ve Větě 3.1.

Inkluze 3 platí, protože $\mathbf{Ax} = \mathbf{0}$ implikuje $\mathbf{BAx} = \mathbf{0}$ (po vynásobení maticí \mathbf{B} zleva). Rovnost 4 platí, neboť $\text{null } \mathbf{B} = \{\mathbf{0}\}$, což je implikace $3 \Rightarrow 1$ ve Větě 3.2. \square

Jako důsledek dostaneme slíbený důkaz nerovnosti (2.4). Z tvrzení 1 Věty 3.4 a rovnosti (3.10) máme $\text{rank}(\mathbf{AB}) \leq \text{rank } \mathbf{A}$. Ovšem díky (2.2) máme také

$$\text{rank}(\mathbf{AB}) = \text{rank}((\mathbf{AB})^T) = \text{rank}(\mathbf{B}^T \mathbf{A}^T) \leq \text{rank}(\mathbf{B}^T) = \text{rank } \mathbf{B}.$$

Dáno dohromady, máme $\text{rank}(\mathbf{AB}) \leq \min\{\text{rank } \mathbf{A}, \text{rank } \mathbf{B}\}$.

Dále uvedeme poměrně překvapivou větu, kterou budeme potřebovat později.

Věta 3.5. *Pro libovolnou matici \mathbf{A} platí*

$$\text{rng}(\mathbf{AA}^T) = \text{rng } \mathbf{A} \tag{3.13a}$$

$$\text{null}(\mathbf{AA}^T) = \text{null } \mathbf{A}^T. \tag{3.13b}$$

Důkaz. Platí $\text{null}(\mathbf{AA}^T) \subseteq \text{null } \mathbf{A}^T$, neboť

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

Zde druhá implikace platí proto, že pro libovolný vektor \mathbf{y} máme $\mathbf{y}^T \mathbf{y} = 0 \Rightarrow \mathbf{y} = \mathbf{0}$ (napište si součin $\mathbf{y}^T \mathbf{y}$ ve složkách!). Ale z Věty 3.4 je také $\text{null}(\mathbf{AA}^T) \supseteq \text{null } \mathbf{A}^T$, tedy platí (3.13b).

Z (3.13b) a (3.12) máme $\dim \text{rng}(\mathbf{AA}^T) = \text{rank}(\mathbf{AA}^T) = \text{rank } \mathbf{A} = \dim \text{rng } \mathbf{A}$. Z Věty 3.4 máme $\text{rng}(\mathbf{AA}^T) \subseteq \text{rng } \mathbf{A}$. Ale pokud je podprostor podmnožinou jiného podprostoru a oba mají stejnou dimenzi, musí být stejné. To je jasné: libovolná báze $\text{rng}(\mathbf{AA}^T)$ leží také v $\text{rng } \mathbf{A}$, a protože oba podprostory mají stejnou dimenzi, je to také báze $\text{rng } \mathbf{A}$. \square

3.3 Afinní podprostor a zobrazení

Afinní kombinace vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je lineární kombinace (3.1), 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. **Afinní podprostor**¹ lineárního prostoru \mathbb{R}^n je množina $A \subseteq \mathbb{R}^n$, která je uzavřená vůči afinním kombinacím (tedy každá afinní kombinace vektorů z A leží v A).

Příklad 3.6. Mějme dva lineárně nezávislé vektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Jejich lineární obal je množina

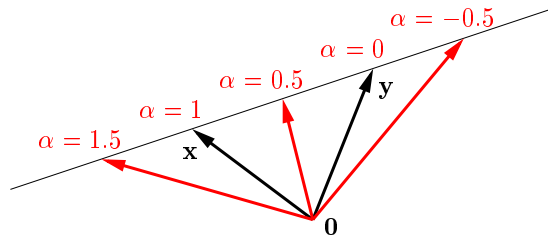
$$\text{span}\{\mathbf{x}, \mathbf{y}\} = \{\alpha \mathbf{x} + \beta \mathbf{y} \mid \alpha, \beta \in \mathbb{R}\},$$

tedy rovina procházející těmito dvěma body a počátkem $\mathbf{0}$, tedy celý \mathbb{R}^2 . Jejich afinní obal je množina

$$\text{aff}\{\mathbf{x}, \mathbf{y}\} = \{\alpha \mathbf{x} + \beta \mathbf{y} \mid \alpha, \beta \in \mathbb{R}, \alpha + \beta = 1\} = \{\alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \mid \alpha \in \mathbb{R}\},$$

což je přímka procházející body \mathbf{x}, \mathbf{y} . Na obrázku jsou vektory $\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}$ pro různé hodnoty α :

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



Podobně, lineární obal dvou lineárně nezávislých vektorů z \mathbb{R}^3 je rovina procházející těmito dvěma body a počátkem $\mathbf{0}$, avšak jejich afinní obal je přímka procházející těmito dvěma body. Afinní obal tří lineárně nezávislých bodů v \mathbb{R}^3 je rovina procházející těmito třemi body. \square

Základní pozorování: afinní kombinace nezávisí na počátku. Proto si prvky vektorového prostoru představujeme jako šipky a prvky afinního prostoru jako body.

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

$$X + \mathbf{y} = \{ \mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X \}. \quad (3.14)$$

Věta 3.6.

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

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

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

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

Věta ukazuje, že afinní podprostor není nic jiného než ‘posunutý’ lineární podprostor (tedy nemusí procházet počátkem, na rozdíl od lineárního podprostoru). **Dimenze afinního podprostoru** je dimenze tohoto lineárního podprostoru. Afinnímu podprostoru \mathbb{R}^n dimenze 0, 1, 2 a $n - 1$ se říká po řadě **bod**, **přímka**, **rovina** a **nadrovina**.

Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ nazveme **afinní**, pokud (3.3) platí pro všechna $\alpha_1 + \dots + \alpha_k = 1$, tedy zobrazení afinní kombinace je rovno afinní kombinaci zobrazení. Lze ukázat (proved’te!), že zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je afinní právě tehdy, když existuje matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a vektor $\mathbf{b} \in \mathbb{R}^m$ tak, že

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}. \quad (3.15)$$

Příklad 3.7. Zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x_1, x_2) = (x_1 + x_2 + 1, x_1 - x_2, 2x_1)$ je afinní. To bychom mohli dokázat ověřením podmínek (3.3) pro $\alpha_1 + \dots + \alpha_k = 1$. Ale je to patrné i z toho, že zobrazení lze vyjádřit ve tvaru (3.15):

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

Pro $m = 1$ je zobrazení (3.15) **afinní funkce**² $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a má tvar

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \cdots + a_n x_n + b, \quad (3.16)$$

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

3.4 Cvičení

Přidat cvičení: Najděte matici daného lineárního zobrazení.

- 3.1. Rozhodněte, zda následující množiny tvoří lineární nebo afinní podprostor \mathbb{R}^n a případně určete jeho dimenzi:
- $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = 0 \}$ pro dané $\mathbf{a} \in \mathbb{R}^n$
 - $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = \alpha \}$ pro dané $\mathbf{a} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$
 - $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{x} = 1 \}$
 - $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \mathbf{x}^T = \mathbf{I} \}$ pro dané $\mathbf{a} \in \mathbb{R}^n$
 - $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 0 \}$
 - $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} = \mathbf{b} \}$ pro dané $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$
- 3.2. Je dáno zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{x} \times \mathbf{y}$, kde $\mathbf{y} \in \mathbb{R}^3$ je pevný vektor a \times označuje vektorový součin. Jde tedy o zobrazení $\mathbb{R}^3 \rightarrow \mathbb{R}^3$. Je toto zobrazení lineární? Pokud ano, najděte matici \mathbf{A} tak, aby $\mathbf{f}(\mathbf{x}) = \mathbf{A} \mathbf{x}$. Čemu je rovno \mathbf{A}^T ? Jakou hodnotu má \mathbf{A} ?
- 3.3. Máme zobrazení $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definované jako $\mathbf{f}(x, y) = (x + y, 2x - 1, x - y)$. Je toto zobrazení lineární? Pokud ano, napište ho ve formě (3.4). Je toto zobrazení afinní? Pokud ano, napište ho ve formě (3.15). Obě odpovědi dokažte z definic.
- 3.4. Dokažte, že
- množina řešení homogenní lineární soustavy $\mathbf{A} \mathbf{x} = \mathbf{0}$ je lineární podprostor,
 - množina řešení nehomogenní lineární soustavy $\mathbf{A} \mathbf{x} = \mathbf{b}$ ($\mathbf{b} \neq \mathbf{0}$) je afinní podprostor.
- 3.5. Najděte prostor obrazů a nulový prostor následujících lineárních zobrazení:
- $\mathbf{f}(x_1, x_2, x_3) = (x_1 - x_2, x_2 - x_3 + 2x_1)$
 - $\mathbf{f}(x_1, x_2) = (2x_1 + x_2, x_1 - x_2, 2x_2 + x_1)$
- 3.6. Máte matice \mathbf{A} a \mathbf{B} se stejným počtem řádků. Jak byste ověřili, zda $\text{rng } \mathbf{A} = \text{rng } \mathbf{B}$, umíte-li spočítat hodnotu libovolné matice?
- 3.7. Které z těchto výroků jsou pravdivé? Každý výrok dokažte nebo najděte protipříklad. Některé výroky mohou platit jen pro určité rozměry matic – najděte tedy co nejobecnější podmínky na rozměry matic, aby výroky byly pravdivé.
- Pokud $\mathbf{A} \mathbf{B}$ má plnou hodnotu, pak \mathbf{A} a \mathbf{B} mají plnou hodnotu.
 - Pokud \mathbf{A} a \mathbf{B} mají plnou hodnotu, pak $\mathbf{A} \mathbf{B}$ má plnou hodnotu.
 - Pokud \mathbf{A} a \mathbf{B} mají triviální nulový prostor, pak $\mathbf{A} \mathbf{B}$ má triviální nulový prostor.

² V lineární algebře znamená slovo ‘lineární’ něco jiného než v matematické analýze. Např. funkci jedné proměnné $f(x) = ax + b$ znáte ze základní školy jako lineární, v lineární algebře však lineární není – je afinní. Ovšem soustavě rovnic $\mathbf{A} \mathbf{x} = \mathbf{b}$ se říká ‘lineární’ i v lineární algebře.

d) (★) Pokud \mathbf{A} a \mathbf{B} jsou úzké s plnou hodnotí a platí $\mathbf{A}^T \mathbf{B} = \mathbf{0}$, pak matice $[\mathbf{A} \ \mathbf{B}]$ je úzká s plnou hodnotí.

e) (★) Pokud matice $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$ má plnou hodnotu, pak \mathbf{A} i \mathbf{B} mají plnou hodnotu.

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

Nápověda ke cvičením

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

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

3.5.a) $\text{rng } \mathbf{A} = \mathbb{R}^2$, $\text{null } \mathbf{A} = \text{span}\{(1, 1, 3)\}$.

3.5.b) $\text{rng } \mathbf{A} = \text{span}\{(2, 1, 1), (1, -1, 2)\}$, $\text{null } \mathbf{A} = \{(0, 0)\}$.

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

Kapitola 4

Ortogonalita

4.1 Standardní skalární součin

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

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

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

Standardní skalární součin indukuje **eukleidovskou normu**¹

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T \mathbf{x}} = (x_1^2 + \cdots + x_n^2)^{1/2},$$

Norma splňuje **trojúhelníkovou nerovnost** $\|\mathbf{x} + \mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$, která snadno plyne z Cauchyovy-Schwarzovy nerovnosti (umocněte a roznásobte). Norma měří *délku* vektoru \mathbf{x} . Úhel φ dvojice vektorů se spočítá jako

$$\cos \varphi = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2}.$$

Eukleidovská norma indukuje **eukleidovskou metriku**

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2,$$

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

*Protože pro $n = 3$ takto definované pojmy délky, úhlu a vzdálenosti dobře modelují prostor, ve kterém žijeme, prostoru \mathbb{R}^n se standardním skalárním součinem se často říká **Eukleidovský prostor**. Jeho prvkům se místo vektory někdy říká body. Slovo vektor užíváme tehdy, chceme-li zdůraznit, že \mathbf{x} patří do vektorového prostoru, např. směr vektoru spojujícího počátek $\mathbf{0}$ s bodem \mathbf{x} . Slovo bod užíváme, chceme-li zdůraznit afinní či metrickou strukturu.*

4.2 Ortogonální vektory

Dva vektory nazveme **ortogonální** (kolmé), pokud $\mathbf{x}^T \mathbf{y} = 0$, což se značí také $\mathbf{x} \perp \mathbf{y}$.

¹ Pro eukleidovskou normu používáme symbol $\|\cdot\|_2$ místo pouhého $\|\cdot\|$, neboť později potkáme i jiné normy.

Vektor nazveme **normalizovaný**, pokud má jednotkovou délku ($\|\mathbf{x}\|_2 = 1 = \mathbf{x}^T \mathbf{x}$). Množinu vektorů $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ nazveme **ortonormální**, pokud každý vektor z této množiny je normalizovaný a každá dvojice vektorů z této množiny je ortogonální, tedy

$$\mathbf{x}_i^T \mathbf{x}_j = \begin{cases} 0 & \text{když } i \neq j, \\ 1 & \text{když } i = j. \end{cases} \quad (4.1)$$

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

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

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

4.3 Ortogonální podprostory

Podprostory X a Y prostoru \mathbb{R}^n jsou **ortogonální**, je-li $\mathbf{x} \perp \mathbf{y}$ pro každé $\mathbf{x} \in X$ a $\mathbf{y} \in Y$. Značíme $X \perp Y$. Pro testování ortogonality podprostorů nemusíme testovat každou dvojici $\mathbf{x} \in X$ a $\mathbf{y} \in Y$, ale postačuje ověřit, zda každý bázový vektor podprostoru X je kolmý na každý bázový vektor podprostoru Y . Opravdu, je-li $\mathbf{x}_1, \dots, \mathbf{x}_k \in X$ a $\mathbf{y}_1, \dots, \mathbf{y}_l \in Y$ a platí-li $\mathbf{x}_i^T \mathbf{y}_j = 0$ pro každé i, j , pak také $(\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k)^T (\beta_1 \mathbf{y}_1 + \dots + \beta_l \mathbf{y}_l) = 0$.

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

$$X^\perp = \{ \mathbf{y} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{y} = 0 \text{ pro všechna } \mathbf{x} \in X \}. \quad (4.2)$$

Je to tedy množina všech vektorů z \mathbb{R}^n takových, že každý je kolmý na každý vektor z X .

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

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

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

To lze napsat ekvivalentně jako $X^\perp = \text{null } \mathbf{A}$, kde $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{bmatrix}$. \square

Některé vlastnosti ortogonálního doplňku:

- $(X^\perp)^\perp = X$. Z toho plyne, že pokud $X^\perp = Y$, pak také $Y^\perp = X$. Tedy podprostory X, Y jsou ortogonálním doplňkem jeden druhého.
- $\dim X + \dim(X^\perp) = n$
- Pro každý vektor $\mathbf{z} \in \mathbb{R}^n$ existuje právě jeden $\mathbf{x} \in X$ a právě jeden $\mathbf{y} \in X^\perp$ tak, že $\mathbf{z} = \mathbf{x} + \mathbf{y}$.

4.4 Čtyři základní podprostory matice

Každá matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ generuje čtyři **základní podprostory**:

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

Tyto podprostory jsou svázány vztahy

$$(\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T), \quad (4.3a)$$

$$(\text{rng } \mathbf{A})^\perp = \text{null}(\mathbf{A}^T). \quad (4.3b)$$

Tyto vztahy plynou přímo z definice ortogonálního doplňku a z rovností (3.7) a (3.8). Stačí si uvědomit, že např. $\text{rng}(\mathbf{A}^T)$ je množina všech lineárních kombinací řádků matice \mathbf{A} a $\text{null } \mathbf{A}$ je množina všech vektorů kolmých na všechny řádky matice \mathbf{A} ,

4.5 Matice s ortonormálními sloupci

Necht' sloupce matice $\mathbf{U} \in \mathbb{R}^{m \times n}$ tvoří ortonormální množinu vektorů. Protože ortonormální vektory jsou lineárně nezávislé, je nutně $m \geq n$. Podmínku ortonormality (4.1) sloupců matice \mathbf{U} lze psát stručně jako

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}_n. \quad (4.4)$$

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

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

Pro $\mathbf{x} = \mathbf{y}$ dostaneme, že se zachovává také eukleidovská norma, $\|\mathbf{f}(\mathbf{x})\|_2 = \|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2$. Tedy zobrazení \mathbf{f} zachovává délky a úhly. Taková zobrazení se nazývají **isometrie**.

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

$$\mathbf{U}^T\mathbf{U} = \mathbf{I} \iff \mathbf{U}^T = \mathbf{U}^{-1} \iff \mathbf{U}\mathbf{U}^T = \mathbf{I}. \quad (4.5)$$

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

Ekvivalence (4.5) říká, že má-li čtvercová matice ortonormální sloupce, má ortonormální i řádky. Dále říká, že inverze takové matice se spočítá jednoduše transpozicí. Čtvercové matici splňující podmínky (4.5) 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 řádky².

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 $\det\mathbf{U}$ může nabývat dvou hodnot:

² To je možná důvod, proč se čtvercové matici s ortonormálními sloupci (tedy i řádky) neříká 'ortonormální', ale 'ortogonální'. Obdélníková matice s ortonormálními sloupci a čtvercová matice s ortogonálními (ne však ortonormálními) sloupci zvláštní jména nemají.

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

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

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

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

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

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

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

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

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

Pamatujte: násobení libovolné matice \mathbf{A} permutační maticí zleva permutuje řádky matice \mathbf{A} . Násobení matice \mathbf{A} permutační maticí zprava permutuje sloupce matice \mathbf{A} . \square

Jak reprezentovat podprostory? Bud' bází nebo bází ortogonálního doplňku. Je výhodné když báze je ortonormální. Příklady: přímka, rovina ve 3-D. Pokud matice $\mathbf{U} = [\mathbf{U}_1 \quad \mathbf{U}_2]$ s ortogonálními sloupci je sestavena ze dvou bloků, zjevně platí $\text{rng } \mathbf{U}_1 \perp \text{rng } \mathbf{U}_2$. Je-li navíc matice \mathbf{U} čtvercová regulární (speciálně, je-li matice \mathbf{U} ortogonální), jsou prostory $\text{rng } \mathbf{U}_1$ a $\text{rng } \mathbf{U}_2$ navzájem ortogonálními doplňky.

4.6 QR rozklad

Matice \mathbf{A} je **horní trojúhelníková**, pokud $a_{ij} = 0$ pro každé $i > j$ (neboli pod hlavní diagonálou jsou nuly) a je **dolní trojúhelníková**, pokud $a_{ij} = 0$ pro každé $i < j$ (neboli nad hlavní diagonálou jsou nuly).

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

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \quad (4.6)$$

kde $\mathbf{Q} \in \mathbb{R}^{m \times n}$ má ortonormální sloupce ($\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$) a $\mathbf{R} \in \mathbb{R}^{n \times n}$ je horní trojúhelníková. Pokud \mathbf{A} má plnou hodnost (tedy n) a navíc požadujeme, aby diagonální prvky \mathbf{R} byly kladné

($r_{ii} > 0$), jsou matice \mathbf{Q} a \mathbf{R} dány jednoznačně. V Matlabu je QR rozklad implementován příkazem³ `[Q,R]=qr(A,0)`.

Podívejme se na prostor obrazů a nulový prostor matice \mathbf{A} :

- Jelikož sloupce \mathbf{Q} jsou lineárně nezávislé, dle tvrzení 4 Věty 3.4 máme $\text{null } \mathbf{A} = \text{null } \mathbf{R}$. Z rovnosti (3.12) pak máme $\text{rank } \mathbf{A} = \text{rank } \mathbf{R}$.
- Pokud \mathbf{A} má lineárně nezávislé sloupce (tedy plnou hodnotu), vztah $\text{rank } \mathbf{A} = \text{rank } \mathbf{R}$ znamená, že \mathbf{R} je regulární. Dle tvrzení 2 Věty 3.4 je tedy $\text{rng } \mathbf{A} = \text{rng } \mathbf{Q}$. Vidíme, že v tomto případě lze QR rozklad chápat jako nalezení ortonormální báze podprostoru $\text{rng } \mathbf{A}$, kde tuto bázi tvoří sloupce \mathbf{Q} .

Numerický příklad QR rozkladu.

Zde doplnit plný QR rozklad, u čehož se využije ortog doplněk.

QR rozklad má mnoho aplikací. Typické je jeho užití na řešení lineárních soustav. Např. řešíme soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$ s regulární čtvercovou maticí \mathbf{A} . Rozložíme $\mathbf{A} = \mathbf{Q}\mathbf{R}$ a vynásobíme soustavu zleva \mathbf{Q}^T , což dá

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b}. \quad (4.7)$$

Toto je ekvivalentní úprava, neboť \mathbf{Q} je regulární. Ale protože je \mathbf{R} trojúhelníková, tato soustava se levně vyřeší zpětnou substitucí.

4.6.1 (★) Gramm-Schmidtova ortonormalizace

Oddelat (★).

Gramm-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 následující. Předpokládejme, že již máme vektory $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$. K vektoru \mathbf{a}_k přičteme takovou lineární kombinaci vektorů $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$, aby výsledek byl na všechny z nich ortogonální. Poté tento vektor normalizujeme. Tedy

$$\mathbf{q}_k := \mathbf{a}_k - \sum_{j=1}^{k-1} r_{jk} \mathbf{q}_j \quad (4.8a)$$

$$\mathbf{q}_k := \frac{\mathbf{q}_k}{\|\mathbf{q}_k\|_2}. \quad (4.8b)$$

Algoritmus postupně provede tuto iteraci pro $k = 1, \dots, n$.

Jak najdeme koeficienty r_{jk} ? Z iterace (4.8) vyplývá, že

$$\mathbf{a}_k = \sum_{j=1}^{k-1} r_{jk} \mathbf{q}_j + r_{kk} \mathbf{q}_k = \sum_{j=1}^k r_{jk} \mathbf{q}_j, \quad (4.9)$$

kde jsme označili normalizační faktor v (4.8b) jako r_{kk} . Vztah (4.9) nám dovoluje spočítat koeficienty r_{jk} z požadavku na ortonormalitu vektorů $\mathbf{q}_1, \dots, \mathbf{q}_k$. Jeho vynásobením vektorem \mathbf{q}_i (pro libovolné i) dostaneme

$$r_{ik} = \mathbf{q}_i^T \mathbf{a}_k. \quad (4.10)$$

³ Podotkněme, že příkaz `[Q,R]=qr(A)` spočítá tzv. *plný QR rozklad*, ve kterém je \mathbf{R} horní trojúhelníková stejné velikosti jako \mathbf{A} a \mathbf{Q} je ortogonální velikosti $m \times m$. Zkoumejte tento příkaz pomocí `help qr!`

Soustavu rovností (4.9) lze napsat v maticovém tvaru $\mathbf{A} = \mathbf{QR}$, kde vektory $\mathbf{a}_1, \dots, \mathbf{a}_n$ jsou sloupce matice \mathbf{A} , vektory $\mathbf{q}_1, \dots, \mathbf{q}_n$ jsou sloupce matice \mathbf{Q} , a matice $\mathbf{R} = [r_{ik}]$ je horní trojúhelníková. Všimněte si dále, že vynásobením rovnosti $\mathbf{A} = \mathbf{QR}$ zleva maticí \mathbf{Q}^T dostaneme $\mathbf{Q}^T \mathbf{A} = \mathbf{R}$, neboť $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$. To souhlasí (dle (2.8)) se vztahem (4.10). *Tohle by chtělo ještě rozvést. Vlastně to vede na jiné odvození G-S algoritmu tak, že máme $\mathbf{A} = \mathbf{QR}$, kde $\mathbf{R} = \mathbf{Q}^T \mathbf{A}$. Pokud chceme, aby \mathbf{R} byla horní trojúhelníková, jdou spočítat její členy rekursivně.*

Vylepšenou formu Gramm-Schmidtovy ortonormalizace lze použít pro počítání QR rozkladu. Vylepšení jednak zmenší zaokrouhlovací chyby a jednak dovolí lineární závislost sloupců \mathbf{A} .

Cvičení: Jak reprezentovat přímky a roviny v \mathbb{R}^3 . Bud' bází a posunutím, nebo ortogonálním doplňkem.

4.7 Cvičení

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

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

- 4.9. Necht' X, Y jsou podprostory \mathbb{R}^n . Definujme $X + Y = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}$. Dokažte:
 - a) $X \subseteq Y \implies X^\perp \supseteq Y^\perp$
 - b) $X + Y = \text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_\ell\}$ kde $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ je báze podprostoru X a $\{\mathbf{y}_1, \dots, \mathbf{y}_\ell\}$ je báze podprostoru Y
 - c) $(\star) (X + Y)^\perp = X^\perp \cap Y^\perp$
 - d) $(X \cap Y)^\perp = X^\perp + Y^\perp$

- 4.10. Pokud $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2$, dokažte, že $(\mathbf{x} + \mathbf{y}) \perp (\mathbf{x} - \mathbf{y})$. V prostoru \mathbb{R}^2 nakreslete vektory $\mathbf{x} + \mathbf{y}$ a $\mathbf{x} - \mathbf{y}$.
- 4.11. Mějme vektory $\mathbf{x}_1 = (1, -1, 0, 2)$, $\mathbf{x}_2 = (1, 1, 1, 0)$, $\mathbf{x}_3 = (-1, -1, 2, 0)$. Ověřte, že tyto vektory jsou vzájemně ortogonální. Najděte vektor \mathbf{x}_4 , který je ortogonální na vektory $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$.
- 4.12. Pro $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 1 & 2 & 0 \end{bmatrix}$ najdi ortonormální báze podprostorů $\text{rng } \mathbf{A}$, $\text{null } \mathbf{A}$, $\text{rng}(\mathbf{A}^T)$, $\text{null}(\mathbf{A}^T)$.
Je možno použít počítač.
- 4.13. Spočtete co nejjednodušším způsobem inverzi matice $\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$.

Nápověda ke cvičením

- 4.1. $\text{span}\{(1, 1, -1)\}$
- 4.3. Báze je $\{(1, 1, 1, -1)/2, (3, -1, -1, 1)/\sqrt{12}, (0, 1, 1, 2)/\sqrt{6}\}$
- 4.5. n sudé.
- 4.7. $\binom{n}{2} = n(n-1)/2$
- 4.9.c) Plyne z (b).
- 4.9.d) Plyne z (c) s použitím $(X^\perp)^\perp = X$.
- 4.11. $\mathbf{x}_4 = (-1, 1, 0, 1)$.

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 tehdy, když $\mathbf{b} \in \text{rng } \mathbf{A}$ (tedy \mathbf{b} je lineární kombinací sloupců \mathbf{A}), což lze psát také jako $\text{rank}[\mathbf{A} \ \mathbf{b}] = \text{rank } \mathbf{A}$ (Frobeniova věta). Pokud je množina řešení soustavy neprázdná, je to afinní podprostor \mathbb{R}^n (viz Cvičení 3.4).

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

- Soustava nemá řešení. To nastane právě tehdy, když $\mathbf{b} \notin \text{rng } \mathbf{A}$. Taková soustava se často 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 často nazývá **podurčená**. V tom případě můžeme chtít z množiny řešení vybrat jediné, což je tématem §5.2.

5.1 Přibližné řešení ve smyslu nejmenších čtverců

Pokud soustava (5.1) nemá řešení, řešme ji přibližně (což můžeme značit $\mathbf{Ax} \approx \mathbf{b}$). Hledejme takové \mathbf{x} , aby eukleidovská norma vektoru $\mathbf{r} = \mathbf{b} - \mathbf{Ax}$ zbytků (neboli *reziduí*) byla co nejmenší. Úloha se nezmění (proč?), když místo eukleidovské normy budeme minimalizovat její čtverec $\|\mathbf{r}\|_2^2 = \mathbf{r}^T \mathbf{r} = r_1^2 + \dots + r_m^2$. Tedy řešíme úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2. \quad (5.2)$$

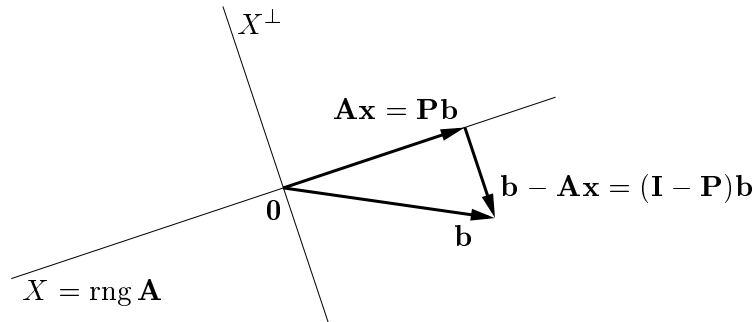
Protože minimalizujeme součet čtverců reziduí, mluvíme o přibližném řešení soustavy **ve smyslu nejmenších čtverců** (*least squares solution*).

Příklad 5.1. Soustava třech rovnic o dvou neznámých

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

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

Úlohu (5.2) vyřešíme následující úvahou. Pokud $\|\mathbf{Ax} - \mathbf{b}\|_2$ (tedy vzdálenost bodů \mathbf{Ax} a \mathbf{b}) má být minimální, musí být vektor $\mathbf{b} - \mathbf{Ax}$ kolmý na prostor $\text{rng } \mathbf{A}$, tedy na každý sloupec matice \mathbf{A} . Obrázek ukazuje situaci:



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

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

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

Soustava (5.3) má řešení pro libovolné \mathbf{A} a \mathbf{b} . To je sice intuitivně zřejmé z naší geometrické úvahy, ovšem není to důkaz. Důkaz plyne z rovnosti (3.13a), neboť $\mathbf{A}^T \mathbf{b} \in \text{rng}(\mathbf{A}^T) = \text{rng}(\mathbf{A}^T \mathbf{A})$, kde $\text{rng}(\mathbf{A}^T \mathbf{A}) = \text{rng}(\mathbf{A}^T)$ je rovnost (3.13a) použitá na matici \mathbf{A}^T .

Dle rovnosti (3.13a) je matice $\mathbf{A}^T \mathbf{A}$ regulární právě tehdy, když matice \mathbf{A} má hodnost n (tedy lineárně nezávislé sloupce). V tom případě můžeme soustavu (5.3) řešit pomocí inverze. Řešením je vektor $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, kde

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T. \quad (5.4)$$

Matice (5.4) se nazývá **pseudoinverze úzké matice \mathbf{A}** . 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 (5.4) nelze použít. V tom případě soustava (5.3), a tedy i úloha (5.2), mají nekonečně mnoho (afinní podprostor) řešení (pozor, to je něco jiného, než že soustava (5.1) má nekonečně mnoho řešení!).

5.1.1 Řešení pomocí QR rozkladu

I když má matice \mathbf{A} lineárně nezávislé sloupce, řešení pomocí pseudoinverze (5.4) nemusí být vhodné pro numerické výpočty, kdy nezbytně používáme aritmetiku s konečnou přesností.

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

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

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

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

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

Numericky vhodnější způsob je řešit normální rovnici *bez* explicitního výpočtu součinu $\mathbf{A}^T \mathbf{A}$. To lze udělat pomocí QR rozkladu $\mathbf{A} = \mathbf{Q} \mathbf{R}$. Po dosazení do normální rovnice máme $\mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} \mathbf{x} = \mathbf{R}^T \mathbf{Q}^T \mathbf{b}$. Po užití $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ a násobení maticí \mathbf{R}^{-T} zleva (což je ekvivalentní operace) máme

$$\mathbf{R} \mathbf{x} = \mathbf{Q}^T \mathbf{b}. \quad (5.5)$$

To je stejný vzorec jako (4.7), rozdíl je jen v tom, že v (4.7) je matice \mathbf{Q} čtvercová a zde obdélníková.

V Matlabu je řešení nehomogenní lineární soustavy implementováno v operátoru `\` (*zpětné lomítko*). Pokud je soustava přeuročená, 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`!

5.1.2 Více o ortogonální projekci

Je poučné dále rozvinout geometrickou úvahu, pomocí níž jsme odvodili normální rovnici. Pokud \mathbf{x} je řešení normální rovnice, vektor $\mathbf{A} \mathbf{x}$ je ortogonální projekcí vektoru \mathbf{b} na podprostor $X = \text{rng } \mathbf{A}$ (viz obrázek výše). Pokud \mathbf{A} má lineárně nezávislé sloupce (tj. tyto sloupce tvoří bázi podprostoru X), z (5.4) máme $\mathbf{A} \mathbf{x} = \mathbf{P} \mathbf{b}$, kde

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

Dostali jsme důležitý výsledek: ortogonální projekce vektoru na podprostor X je lineární zobrazení reprezentované maticí \mathbf{P} . Proto se tato matice často nazývá **projektor**.

Pokud X je reprezentován ortonormální bází, tedy $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, výraz (5.6) se zjednoduší na $\mathbf{P} = \mathbf{A} \mathbf{A}^T$. Připomeňme (viz §4.5), že obdélníková matice \mathbf{A} s ortonormálními sloupci nemusí mít ortonormální řádky, neboli $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ neimplikuje $\mathbf{A} \mathbf{A}^T = \mathbf{I}$. Nabízela se otázka, co je tedy matice $\mathbf{A} \mathbf{A}^T$. Zde jste dostali odpověď.

Speciálním případem ortogonální projekce je případ $\dim X = 1$, tedy projekce na přímku. Necht' $X = \text{span}\{\mathbf{a}\}$, kde předpokládáme $\|\mathbf{a}\|_2 = 1$. Pak $\mathbf{P} = \mathbf{a} \mathbf{a}^T$. Vzoreček $(\mathbf{a}^T \mathbf{b}) \mathbf{a} = \mathbf{a} \mathbf{a}^T \mathbf{b} = \mathbf{P} \mathbf{b}$ pro průmět vektoru \mathbf{b} na normalizovaný vektor \mathbf{a} máte znát ze střední školy.

Čistě geometrickou úvahou vidíme, co je prostorem obrazů a nulovým prostorem projektoru. Libovolný vektor z \mathbb{R}^m se promítne na podprostor X . Libovolný vektor kolmý na X se promítne do nulového vektoru $\mathbf{0}$. Tedy

$$\begin{aligned} \text{rng } \mathbf{P} &= X \\ \text{null } \mathbf{P} &= X^\perp. \end{aligned}$$

Z obrázku je dále patrné, že vektor $\mathbf{b} - \mathbf{A} \mathbf{x} = \mathbf{b} - \mathbf{P} \mathbf{b} = (\mathbf{I} - \mathbf{P}) \mathbf{b}$ je ortogonální projekcí vektoru \mathbf{b} na X^\perp . Tedy projektor na X^\perp je matice $\mathbf{I} - \mathbf{P}$. Všimněte si, že projektor na X^\perp má přirozenou úlohu při přibližném řešení soustavy: hodnota minima problému (5.2) je $\|\mathbf{b} - \mathbf{A} \mathbf{x}\|_2^2 = \|\mathbf{b} - \mathbf{P} \mathbf{b}\|_2^2 = \|(\mathbf{I} - \mathbf{P}) \mathbf{b}\|_2^2$.

Poznámka o obecné projekci. *Projekcí* se v lineární algebře rozumí každé lineární zobrazení $f(\mathbf{y}) = \mathbf{P}\mathbf{y}$, které splňuje $f(f(\mathbf{y})) = f(\mathbf{y})$, tedy $\mathbf{P}\mathbf{P} = \mathbf{P}^2 = \mathbf{P}$. To vyjadřuje pochopitelný požadavek, že když jednou vektor promítneme, tak další promítnutí ho již nezmění. Projekce obecně nemusí být ortogonální, může být šikmá – pak promítáme podle podprostoru $\text{null } \mathbf{P}$ na podprostor $\text{rng } \mathbf{P}$. Projekce je ortogonální, když¹ $\text{null } \mathbf{P} \perp \text{rng } \mathbf{P}$. To nastane právě tehdy, když kromě $\mathbf{P}^2 = \mathbf{P}$ platí navíc $\mathbf{P}^T = \mathbf{P}$ (důkaz tohoto tvrzení vynecháme). Zkontrolujte si, že projektor definovaný vzorcem (5.6) splňuje $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^T$!

5.2 Řešení s nejmenší normou

Předpokládejme nyní, že soustava (5.1) je podurčená, neboli má nekonečně mnoho řešení. Necht' \mathbf{x}' je libovolný vektor splňující $\mathbf{A}\mathbf{x}' = \mathbf{b}$ (tzv. **partikulární řešení** soustavy). Protože pro každé $\mathbf{x} \in \text{null } \mathbf{A}$ je $\mathbf{A}(\mathbf{x}' + \mathbf{x}) = \mathbf{A}\mathbf{x}' = \mathbf{b}$, množinu řešení soustavy lze psát parametricky jako (viz §3.3)

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\} = \mathbf{x}' + \text{null } \mathbf{A}. \quad (5.7)$$

Je často užitečné z této množiny řešení vybrat jediné podle nějakého kritéria. Přírodním kritériem je minimalizovat euklidovskou normu (tedy vzdálenost od počátku) řešení, což vede na úlohu

$$\min\{\|\mathbf{x}\|_2^2 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}. \quad (5.8)$$

Místo normy $\|\mathbf{x}\|_2$ opět minimalizujeme její čtverec. Tato úloha je známa jako řešení nehomogenní lineární soustavy **s nejmenší normou** (*least norm solution*). Podotkněme, že někdy je vhodné použít jiná kritéria než nejmenší eukleidovskou normu, viz např. Cvičení 9.25.

Příklad 5.3. Soustava dvou rovnic o třech neznámých

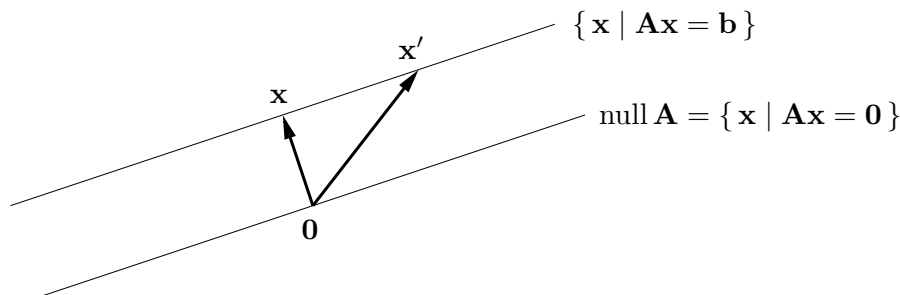
$$\begin{aligned} x + 2y + z &= 1 \\ -x + y + 2z &= 2 \end{aligned}$$

je podurčená, tj. má nekonečně mnoho řešení. Množina řešení je

$$(x_0, y_0, z_0) + \text{null } \mathbf{A} = (1, -1, 2) + \text{span}\{(1, -1, 1)\} = \{(1 + \alpha, -1 - \alpha, 2 + \alpha) \mid \alpha \in \mathbb{R}\}.$$

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

Úlohu (5.8) by bylo vhodné řešit metodou Lagrangeových multiplikátorů, což se ale naučíme až v pozdější kapitole. Nyní ji vyřešíme geometrickou úvahou. Tato úvaha bude naopak rozvíčkou pro pozdější odvození metody Lagrangeových multiplikátorů v 9.4.1.



¹ Pro obecnou čtvercovou matici samozřejmě neplatí, že její nulový prostor a prostor obrazů jsou navzájem ortogonální, tím méně ortogonálním doplňkem. Neplést se vztahy (4.3)!

Vektory \mathbf{x} a \mathbf{x}' jsou dvě různá řešení soustavy, ale pouze \mathbf{x} má nejmenší normu. Z obrázku je vidět, že řešení \mathbf{x} má nejmenší normu právě tehdy, když vektor \mathbf{x} je kolmý na nulový prostor matice \mathbf{A} . Tedy $\mathbf{x} \in (\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T)$, kde poslední rovnost je (4.3a). To ale znamená, že \mathbf{x} musí být lineární kombinací řádků \mathbf{A} . Neboli musí existovat vektor $\boldsymbol{\lambda} \in \mathbb{R}^m$ tak, že $\mathbf{x} = \mathbf{A}^T \boldsymbol{\lambda}$. Pro vyřešení úlohy (5.8) tedy musíme vyřešit soustavu rovnic

$$\mathbf{A}^T \boldsymbol{\lambda} = \mathbf{x}, \quad (5.9a)$$

$$\mathbf{A} \mathbf{x} = \mathbf{b}. \quad (5.9b)$$

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

Vyřešme tuto soustavu. Dosadíme \mathbf{x} do druhé rovnice, $\mathbf{A} \mathbf{A}^T \boldsymbol{\lambda} = \mathbf{b}$. Předpokládejme, že matice $\mathbf{A} \mathbf{A}^T$ má plnou hodnotu, což dle rovnosti (3.13a) nastane právě tehdy, když matice \mathbf{A} má hodnotu m (tedy lineárně nezávislé řádky). Potom $\boldsymbol{\lambda} = (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{b}$. Dosazením do první rovnice dostaneme $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, kde

$$\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}. \quad (5.10)$$

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

5.2.1 Pseudoinverze obecné matice s plnou hodnotou

Pseudoinverzi úzké matice jsme definovali dříve vzorcem (5.4). Nyní shrňme: má-li matice \mathbf{A} plnou hodnotu (tedy $\min\{m, n\}$), definujeme její pseudoinverzi jako

$$\mathbf{A}^+ = \begin{cases} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T & \text{když } m \geq n, \\ \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} & \text{když } m \leq n. \end{cases} \quad (5.11)$$

Vektor $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$ je v prvním případě řešení soustavy $\mathbf{A} \mathbf{x} = \mathbf{b}$ ve smyslu nejmenších čtverců, ve druhém případě řešení soustavy s nejmenší normou. Pokud $m = n$, je v obou případech $\mathbf{A}^+ = \mathbf{A}^{-1}$ (ověřte!).

V případě, že \mathbf{A} nemá plnou hodnotu, vzorec (5.11) nelze použít a pseudoinverzi definujeme jinak. K tomu se vrátíme později v §7.6.

5.3 Použití

Ve tvaru (5.2) lze zformulovat mnoho užitečných úloh. To často není na první pohled vidět a proto s tím studenti mívají problémy.

Příklad 5.4. Hledejme příčku (nejkratší spojnicí) dvou mimoběžných přímek v prostoru \mathbb{R}^n . Necht' i -tá přímka je zadána dvěma body, které na ní leží, označme je $\mathbf{p}_i, \mathbf{q}_i \in \mathbb{R}^n$ pro $i = 1, 2$. Chceme tuto úlohu zformulovat ve tvaru (5.2). Řešíme přeúřčenou soustavu

$$\mathbf{p}_1 + t_1(\mathbf{q}_1 - \mathbf{p}_1) \approx \mathbf{p}_2 + t_2(\mathbf{q}_2 - \mathbf{p}_2).$$

Tato soustava má n rovnic a 2 neznámé t_1, t_2 . Jde napsat jako $\mathbf{A} \mathbf{x} \approx \mathbf{b}$ kde

$$\mathbf{A} = [\mathbf{q}_1 - \mathbf{p}_1 \quad \mathbf{p}_2 - \mathbf{q}_2], \quad \mathbf{x} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \quad \mathbf{b} = \mathbf{p}_2 - \mathbf{p}_1. \quad \square$$

5.3.1 Lineární regrese

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

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

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

$$\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^m (y_i - f(t_i, \mathbf{x}))^2. \quad (5.12)$$

Zvolme regresní funkci jako lineární kombinaci

$$f(t, \mathbf{x}) = x_1 \varphi_1(t) + \dots + x_n \varphi_n(t) = \boldsymbol{\varphi}(t)^T \mathbf{x} \quad (5.13)$$

daných funkcí² $\varphi_i: T \rightarrow \mathbb{R}$. Pak

$$\sum_{i=1}^m (y_i - f(t_i, \mathbf{x}))^2 = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2,$$

kde $\mathbf{y} = (y_1, \dots, y_m)$ a prvky matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ jsou $a_{ij} = \varphi_j(t_i)$ (odvod'te!). Tedy vyjádřili jsme úlohu (5.12) ve tvaru (5.2). Protože funkce (5.13) je lineární funkcí parametrů \mathbf{x} , mluvíme o **lineární regresi**.

Příklad 5.5. *Polynomiální regrese*³. Necht' $T = \mathbb{R}$ a $\varphi_i(t) = t^{i-1}$. Pak regresní funkce je polynom stupně $n - 1$,

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

Matice

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

je známá jako *Vandermondova matice*.

Speciálně pro $n = 1$ úloha (5.12) je $\min_x \sum_i (y_i - x)^2$. Řešením je aritmetický průměr $x = \frac{1}{m} \sum_{i=1}^m y_i$ (ověřte!). \square

Regularizace: Minimalizujeme $\|\mathbf{A}\mathbf{x} - \mathbf{b}\| + \mu\|\mathbf{x}\|$. Jde to převést na nejmenší čtverce. Pro $\mu \rightarrow 0$ dostaneme least norm řešení.

Zde by měla následovat celá jedna přednáška o aplikacích LS a LN. To jsem zkusil v roce 2-13, ale moc se mi to nepovedlo. Říkal jsem tyto příklady: příčka mimoběžek; regrese, její použití, křížová validace; identifikace jednoduchého lineárního dynamického systému dle Boyda slides 6-11, moving average, autonomní systém; posunutí tělesa do daného stavu s minimálním součtem sil dle Boyda slides 8-10, Pareto minimum a regularizace (ale to se mi moc nepovedlo). Nestihl jsem rekonstrukci signálu s vyhlazením příp. rekonstrukci z CT, lineární zobrazení kolem nás (most jako elastická struktura, příklady z Boyda).

² Funkce φ_i se často nazývají *bázové funkce*, a to i tehdy (což je protimluv), když jsou lineárně závislé.

³ Nedejte se zmást tím, že polynom není lineární funkce a přesto jde o lineární regresi. Důležité je, že regresní funkce (5.13) je lineární v parametrech \mathbf{x} .

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? Kladné i záporné odpovědi dokažte.

a) Pokud $m < n$, pak soustava má vždy řešení.

b) Pokud $m > n$, pak soustava nemá nikdy řešení.

c) 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. Hledáme bod $\mathbf{x} \in \mathbb{R}^m$, který minimalizuje součet čtverců vzdáleností k daným bodům $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$, tj. výraz $\sum_{i=1}^n \|\mathbf{a}_i - \mathbf{x}\|_2^2$. Zformulujte úlohu ve tvaru (5.2) (analogicky Příkladu 5.4). Dokažte, že minimum se nabývá v těžišti $\mathbf{x} = \frac{1}{n} \sum_{i=1}^n \mathbf{a}_i$.

5.4. Jsou dány vektory $\mathbf{a}, \mathbf{s}, \mathbf{y} \in \mathbb{R}^n$. Hledáme vzdálenost bodu \mathbf{y} od přímky $\{\mathbf{a} + t\mathbf{s} \mid t \in \mathbb{R}\}$. Zformulujte ve tvaru (5.2).

5.5. Jsou dány vektory $\mathbf{a}, \mathbf{y} \in \mathbb{R}^n$ a skalár $b \in \mathbb{R}$. Hledáme vzdálenost bodu \mathbf{y} od nadroviny $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$. Zformulujte ve tvaru (5.2).

5.6. Máme m přímek v rovině, přičemž i -tá přímka má rovnici $\mathbf{a}_i^T \mathbf{x} = b_i$ pro dané $\mathbf{a}_i \in \mathbb{R}^2$ a $b_i \in \mathbb{R}$. Hledáme bod, který minimalizuje součet čtverců vzdáleností od jednotlivých přímek. Zformulujte ve tvaru (5.2).

5.7. 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 odhad souřadnic x_1, \dots, x_n . Zformulujte ve tvaru (5.2). Lze dosáhnout, aby měla matice \mathbf{A} plnou hodnotu? Pokud ne, jak byste úlohu změnil, aby měla?

5.8. 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 (nápopověda: sdužte čísla w_i do diagonální matice $\mathbf{W} = \text{diag}(\mathbf{w})$). Napište maticový výraz pro optimální \mathbf{x} . Za jakých podmínek má úloha řešení?

5.9. 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) $(\text{span}\{\mathbf{u}, \mathbf{v}\})^\perp$.

5.10. 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 . Nápopověda: Nejsou náhodou vektory ortonormální?

- 5.11. 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) $\text{null}(\mathbf{A}^T)$.
- 5.12. 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 (5.6) a podprostor $X = \text{rng } \mathbf{A}$? Jaký je geometrický význam této situace?
- 5.13. (★) V §5.1.2 jsme ukázali, že matice $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ má význam projekce na podprostor $\text{rng } \mathbf{A}$. Na základě úvah v §5.2 je přirozené vytvořit podobnou matici $\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}$. Jaký geometrický význam má tato matice?
- 5.14. (★) Pro $\|\mathbf{a}\|_2 = 1$ je $\mathbf{H} = \mathbf{I} - 2\mathbf{a}\mathbf{a}^T$ známa jako *Householderova matice*. Transformace $\mathbf{H}\mathbf{x}$ je zrcadlení vektoru \mathbf{x} kolem nadroviny s normálovým vektorem \mathbf{a} , proto se \mathbf{H} také někdy nazývá *elementární reflektor*.
- Odvoďte matici \mathbf{H} podobnou úvahou, jako jsme odvodili projektor.
 - Ukažte, že $\mathbf{H} = \mathbf{H}^T$ a $\mathbf{H}^T\mathbf{H} = \mathbf{I}$ (tj. matice je symetrická a ortogonální).
 - Z těchto dvou vlastností vyplývá $\mathbf{H}\mathbf{H} = \mathbf{I}$. Co to říká o transformaci $\mathbf{H}\mathbf{x}$?
 - Ukažte, že $\det \mathbf{H} = -1$.
 - Co je $\mathbf{H}\mathbf{a}$? Co je $\mathbf{H}\mathbf{x}$, když $\mathbf{a}^T\mathbf{x} = 0$? Ukažte algebraicky a odůvodněte geometricky.
- 5.15. V §5.1.2 jsme definovali projekci jako matici \mathbf{P} splňující $\mathbf{P}^2 = \mathbf{P}$. Geometrickou úvahou (tedy bez počítání) najdete aspoň jedno vlastní číslo a příslušný vlastní vektor projekce.
- 5.16. (★) Geometrickou úvahou najdete aspoň dva vlastní vektory a příslušná vlastní čísla Householderovy matice ze Cvičení 5.14.
- 5.17. (★) *RQ rozklad* rozloží matici $\mathbf{A} = \mathbf{R}\mathbf{Q}$, kde \mathbf{R} je horní trojúhelníková a \mathbf{Q} je ortogonální. Jak byste spočítali RQ rozklad z QR rozkladu?
- 5.18. (★) Matice \mathbf{A} se nazývá *normální*, pokud $\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T$. Příkladem je symetrické matice (ale ne každá normální matice je symetrická). Dokažte, že pro normální matice platí $\text{rng } \mathbf{A} \perp \text{null } \mathbf{A}$. Nápoředa: Vycházejte z rovnosti (3.13a).
- 5.19. Dokažte následující vlastnosti pseudoinverze ze vztahu (5.11) pro libovolné matice plné hodnosti:
- $\mathbf{A}^+ = \mathbf{A}^{-1}$ když \mathbf{A} je čtvercová
 - $(\mathbf{A}^+)^+ = \mathbf{A}$
 - $(\mathbf{A}^T)^+ = (\mathbf{A}^+)^T$
 - $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$, $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$, $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$, $(\mathbf{A}^+\mathbf{A})^T = \mathbf{A}^+\mathbf{A}$
 - $\mathbf{A}^T = \mathbf{A}^T\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+\mathbf{A}\mathbf{A}^T$
 - $(\mathbf{A}^T\mathbf{A})^+ = \mathbf{A}^+(\mathbf{A}^T)^+$, $(\mathbf{A}\mathbf{A}^T)^+ = (\mathbf{A}^T)^+\mathbf{A}^+$
- 5.20. Uvažujme projektor (5.6). Báze podprostoru X , na který promítáme, je tvořena sloupci matice \mathbf{A} . Projektor \mathbf{P} se nesmí změnit, vezmeme-li jinou bázi podprostoru. Různé báze podprostoru jsou dány sloupci matice $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{C}$ pro různé regulární matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ (tedy \mathbf{C} je matice přechodu k jiné bázi). Ověřte, že vskutku $\tilde{\mathbf{A}}(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$.

Nápověda ke cvičením

5.2. $(x_1, x_2, x_3) = (2, 1, 0)/3$

5.9. (a) $(2, 1, -3)/14$, (b) $(26, -1, 17)/14$, (c) $(62, -35, 17)/38$, (d) $(14, 35, 21)/38$

5.10. Projektor na X je $\mathbf{P} = \text{diag}(1, 0, 1, 1)$. Projektor na X^\perp je $\mathbf{P} = \text{diag}(0, 1, 0, 0)$.

5.11. (a) $(1, 1, 1)$, (b) $(0.4, -0.2, 0)$, (c) $(0.6, 1.2, 1)$, (d) $(0, 0, 0)$

Kapitola 6

Spektrální rozklad a kvadratické funkce

Nejdříve si řekneme něco o polynomech. Ze základní školy znáte polynomy jedné proměnné, zde nás ale budou zajímat polynomy více proměnných.

Monom k -tého stupně je funkce $a_{k_1 \dots k_n} x_1^{k_1} \dots x_n^{k_n}$, kde $a_{k_1 \dots k_n} \in \mathbb{R}$, $k_1, \dots, k_n \in \{0, 1, \dots, m\}$, a $k_1 + \dots + k_n = k$. **Polynom** n proměnných je součet **monomů**. **Stupeň polynomu** je stupeň jeho monomu nejvyššího stupně. Např. funkce

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

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

Vidíme, že afinní funkce na \mathbb{R}^n (3.16) je jen jiný název pro polynom prvního stupně a lineární funkce (3.5) je jiný název pro homogenní polynom prvního stupně. Podobně, polynom druhého stupně na \mathbb{R}^n se nazývá *kvadratická funkce* a homogenní polynom druhého stupně *kvadratická forma*. V této kapitole se budeme zabývat extrémy kvadratických forem a funkcí.

6.1 Vlastní čísla a vektory

Necht' pro čtvercovou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$, vektor $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq \mathbf{0}$ a skalár $\lambda \in \mathbb{C}$ platí

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

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

Geometrická interpretace.

Rovnici (6.2) lze přepsat jako

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

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

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

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

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

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

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

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

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

Z definice determinantu (2.6) plyne (promyslete!), že charakteristický polynom má stupeň n . Podle *základní věty algebry* má tedy n komplexních kořenů, z nichž některé mohou být násobné. Označíme-li kořeny $\lambda_1, \dots, \lambda_n$, platí

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

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

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

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

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

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

Jak se počítají vlastní čísla a vektory? Charakteristický polynom je hlavně teoretický nástroj a přímé hledání jeho kořenů není vhodné pro numerický výpočet¹. Pro větší matice se používají numerické iterační algoritmy, přičemž pro matice různého typu jsou vhodné různé algoritmy. Matlabská funkce `[V,D]=eig(A)` spočítá matice \mathbf{V} a \mathbf{D} splňující (6.5).

6.1.1 Spektrální rozklad

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

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

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

Je známo mnoho vlastností matic, které postačují pro diagonalizovatelnost. Nejznámější z nich je symetrie.

¹Naopak, hledání kořenů libovolného polynomu lze převést na hledání vlastních čísel matice, která se nazývá *doprovodná matice (companion matrix)* polynomu.

Věta 6.1. Necht' matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická. Pak všechna její vlastní čísla jsou reálná a existuje ortonormální množina n jejích vlastních vektorů.

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

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^T. \quad (6.7)$$

Vlastní čísla a vektory jsou rozsáhlé téma, které jsme zde zdaleka nevyčerpali. Dále nás bude zajímat pouze spektrální rozklad symetrické matice.

6.2 Kvadratická forma

Kvadratická forma je polynom $f: \mathbb{R}^n \rightarrow \mathbb{R}$ daná vzorcem

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (6.8)$$

pro nějakou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$.

Příklad 6.2. Příkladem kvadratické formy dvou proměnných je funkce

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

Každou čtvercovou matici můžeme psát jako součet symetrické a antisymetrické matice:

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

(viz Cvičení 2.11). Ale

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

nebot' $\mathbf{x}^T (\mathbf{A} - \mathbf{A}^T) \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{A}^T \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} - (\mathbf{x}^T \mathbf{A} \mathbf{x})^T = 0$, kde jsme použili skutečnost, že transpozice skaláru je tentýž skalár.

Tedy když \mathbf{A} není symetrická, můžeme ji nahradit její symetrickou částí $\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$ a kvadratická forma se nezmění. Dále proto budeme předpokládat, že \mathbf{A} je symetrická.

Definice 6.1. Symetrická matice \mathbf{A} je

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

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

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

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

- Je-li \mathbf{A} pozitivně [negativně] semidefinitní, pak je v počátku minimum [maximum].
- Je-li \mathbf{A} pozitivně [negativně] definitní, pak je v počátku ostré minimum [maximum].
- Je-li \mathbf{A} indefinitní, pak kvadratická forma extrém nemá.

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

Věta 6.2. *Symetrická matice je*

- pozitivně [negativně] semidefinitní, právě když má všechna vlastní čísla nezáporná [nekladná]
- pozitivně [negativně] definitní, právě když má všechna vlastní čísla kladná [záporná]
- indefinitní, právě když má alespoň jedno kladné a alespoň jedno záporné vlastní číslo.

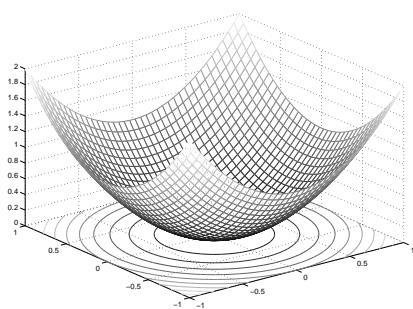
Důkaz. Z rozkladu podle vlastních čísel (6.7) máme

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{V} \mathbf{D} \mathbf{V}^T \mathbf{x} = \mathbf{y}^T \mathbf{D} \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2, \quad (6.9)$$

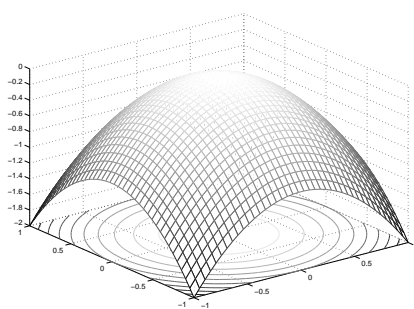
kde $\mathbf{y} = \mathbf{V}^T \mathbf{x}$. Substitutece $\mathbf{x} = \mathbf{V} \mathbf{y}$ tedy diagonalizovala matici kvadratické formy. Protože \mathbf{V} je regulární, definitnost matice \mathbf{A} je stejná jako definitnost matice \mathbf{D} . Ale protože \mathbf{D} je diagonální, její definitnost je okamžitě patrná ze znamének čísel λ_i . Např. výraz (6.9) je nezáporný pro každé \mathbf{y} právě tehdy, když všechna λ_i jsou nezáporná. \square

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

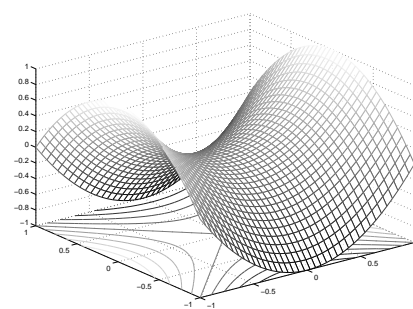
Pro případ dvou proměnných ($n = 2$) si grafy diagonální formy g snadno představíme. Jsou-li obě vlastní čísla kladná, funkce g vypadá jako ‘dolík’. Jsou-li obě vlastní čísla záporná, funkce g vypadá jako ‘kopec’. Mají-li vlastní čísla opačná znaménka, tvarem je ‘sedlo’:



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



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



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

6.3 Kvadratická funkce

Obecnou kvadratickou funkci n proměnných lze psát v maticovém tvaru

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c, \quad (6.10)$$

kde $\mathbf{A}^T = \mathbf{A} \neq \mathbf{0}$. Oproti kvadratické formě tedy přibily lineární a konstantní členy. Naopak, kvadratická forma je kvadratická funkce bez lineárních a konstantních členů. Všimněte si, že pro $n = 1$ je (6.10) známá kvadratická funkce jedné proměnné $f(x) = ax^2 + bx + c$.

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

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

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{A} (\mathbf{x} - \mathbf{x}_0) + y_0. \quad (6.11)$$

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

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

Porovnáním členů stejného stupně máme

$$\mathbf{b} = -2\mathbf{A}\mathbf{x}_0, \quad (6.12a)$$

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

z čehož spočítáme \mathbf{x}_0 a y_0 . Pokud soustava (6.12a) není řešitelná, doplňení na čtverec není možné.

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

Příklad 6.3. Máme kvadratickou funkci

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

Její doplňení na čtverec je

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

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

Příklad 6.4. 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. \square

Příklad 6.5. Řešme znovu úlohu (5.2). Hledáme minimum kvadratické funkce

$$\begin{aligned}
 \|\mathbf{Ax} - \mathbf{b}\|_2^2 &= (\mathbf{Ax} - \mathbf{b})^T (\mathbf{Ax} - \mathbf{b}) \\
 &= (\mathbf{x}^T \mathbf{A}^T - \mathbf{b}^T) (\mathbf{Ax} - \mathbf{b}) \\
 &= \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b} \\
 &= \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - 2\mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b},
 \end{aligned} \tag{6.13}$$

kde jsme použili skutečnost, že skalár je roven své transpozici a tedy $\mathbf{b}^T \mathbf{Ax} = (\mathbf{b}^T \mathbf{Ax})^T = \mathbf{x}^T \mathbf{A}^T \mathbf{b}$. Extrém této kvadratické funkce můžeme najít doplněním na čtverec (viz §6.3). Soustava (6.12a) bude mít tvar $\mathbf{A}^T \mathbf{Ax}_0 = \mathbf{A}^T \mathbf{b}$ (pozor: \mathbf{A} , \mathbf{b} znamená něco jiného v (6.13) a v (6.12a)), 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

$$\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} = (\mathbf{Ax})^T \mathbf{Ax} = \|\mathbf{Ax}\|_2^2 \geq 0. \tag{6.14}$$

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

Kvadrík. Vrstevnice kvadratické funkce se nazývá **kvadrík** (*quadric, quadric surface*). Tedy kvadrík je množina

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{x} + c = 0\}. \tag{6.15}$$

Pro $n = 2$ se kvadrík nazývá **kuželosečka** (*conic*). Důležitým speciálním případem kvadríky je **elipsoidní povrch**², což je množina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{Ax} = 1\}$ pro \mathbf{A} pozitivně definitní.

Zařadit cvičení o převodu kvadratické formy a funkce na maticový tvar a naopak.

6.4 Cvičení

- 6.1. Spočítejte vlastní čísla a vlastní vektory matic $\begin{bmatrix} 1 & 2 \\ -1 & -3 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$.
- 6.2. Napište rovnici, jejímiž kořeny jsou vlastní čísla matice $\begin{bmatrix} 2 & 0 & 3 \\ 0 & -2 & -1 \\ 3 & -1 & 2 \end{bmatrix}$.
- 6.3. Najděte všechna vlastní čísla a vlastní vektory (a) nulové, (b) jednotkové, (c) diagonální matice. Najděte vlastní čísla trojúhelníkové matice.
- 6.4. Ukažte, že $\lambda_1 \times \dots \times \lambda_n = \det \mathbf{A}$. Ukažte (a) pro libovolnou diagonalizovatelnou matici, (b) (*) pro libovolnou čtvercovou matici.
- 6.5. Známe vlastní čísla a vektory matice \mathbf{A} . Jaké jsou vlastní čísla a vektory matice $\mathbf{A} + \alpha \mathbf{I}$?
- 6.6. Je dána kvadratická funkce $f(x, y) = 2x^2 + 2xy - 3y^2 - x + 2y$. Napište ji ve tvaru (6.10) (se symetrickou \mathbf{A}).
- 6.7. Je známo, že libovolnou rotaci ve třírozměrném prostoru lze realizovat jako rotaci kolem jisté přímky (jdoucí počátkem) o jistý úhel. Geometrickou úvahou (tj. bez počítání) zjistěte co nejvíce o vlastních číslech a vektorech rotační matice rozměru 3×3 .

² Někdy se nazývá též *elipsoid*, ale názvosloví není jednoznačné a někteří autoři elipsoidem rozumí množinu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{Ax} \leq 1\}$. Rozdíl je stejný jako mezi sférou a koulí.

- 6.8. Ukažte, že je-li $\mathbf{A} = \mathbf{VDV}^T$ spektrální rozklad symetrické matice \mathbf{A} , platí $\mathbf{A}^n = \mathbf{VD}^n\mathbf{V}^T$.
- 6.9. Ukažte, že vlastní čísla antisymetrické matice jsou nula nebo čistě imaginární.
- 6.10. (★) Ukažte, že dvě čtvercové matice komutují, právě když mají stejné vlastní vektory.
- 6.11. (★) Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{B} \in \mathbb{R}^{n \times m}$. Ukažte, že nenulová vlastní čísla matic \mathbf{AB} a \mathbf{BA} jsou stejná.
- 6.12. Pro dané matice určete, zda jsou pozitivně/negativně (semi)definitní nebo indefinitní:

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

- 6.13. Zjistěte, zda mají následující kvadratické funkce extrém a pokud ano, extrém najděte a určete jeho druh (minimum nebo maximum). Použijte doplnění na úplný čtverec.

a) $f(x, y) = x^2 + 4xy - 2y^2 + 3x - 6y + 5$

b) $f(\mathbf{x}) = \mathbf{x}^T \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \mathbf{x} + [2 \quad -1] \mathbf{x}$

- 6.14. Mějme matici $\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 2 & -4 \end{bmatrix}$. Která z následujících tvrzení jsou pravdivá?

a) Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nezáporný pro každé $\mathbf{x} \in \mathbb{R}^2$.

b) Výraz $\mathbf{x}^T \mathbf{A} \mathbf{x}$ je nekladný pro každé $\mathbf{x} \in \mathbb{R}^2$.

c) Funkce $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ má v bodě $\mathbf{x} = \mathbf{0}$ extrém.

- 6.15. (★) Napište v Matlabu funkci `ellipse(Q)`, která vykreslí elipsu s rovnicí $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$ pro pozitivně definitní \mathbf{A} . Zamyslete se, jak byste postupovali při návrhu funkce `conic(Q)`, která vykreslí kuželosečku $\mathbf{x}^T \mathbf{A} \mathbf{x} = 1$ pro \mathbf{A} libovolné definitnosti (nezapomeňte, že obecná kuželosečka může být neomezená, tedy je nutno ji oříznout do daného obdélníku).

- 6.16. Dokažte, že matice $\mathbf{A}^T \mathbf{A}$ je pozitivně semidefinitní pro každou matici \mathbf{A} .

- 6.17. Dokažte, že matice $\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}$ je pozitivně definitní pro každou matici \mathbf{A} a každé $\mu > 0$.

- 6.18. Dokažte, že (čtvercová symetrická) matice je pozitivně definitní právě tehdy, když její inverze je pozitivně definitní.

- 6.19. Musí mít pozitivně semidefinitní matice na diagonále nezáporné prvky? Odpověď dokažte.

- 6.20. (★) Pozitivně semidefinitní matice lze vnímat jako zobecnění nezáporných čísel. Proto se někdy pozitivní semidefinitnost značí $\mathbf{A} \succeq 0$ a pozitivní definitnost $\mathbf{A} \succ 0$. Zápis $\mathbf{A} \succeq \mathbf{B}$ je pak zkratkou $\mathbf{A} - \mathbf{B} \succeq 0$. Na základě této analogie bychom očekávali, že:

a) Pokud $\mathbf{A} \succeq \mathbf{B}$ a $\mathbf{C} \succeq \mathbf{D}$, potom $\mathbf{A} + \mathbf{C} \succeq \mathbf{B} + \mathbf{D}$.

b) Pokud $\mathbf{A} \succeq 0$ a $\alpha \geq 0$, potom $\alpha \mathbf{A} \succeq 0$.

c) Pokud $\mathbf{A} \succeq 0$, potom $\mathbf{A}^2 \succeq 0$.

d) Pokud $\mathbf{A} \succeq 0$ a $\mathbf{B} \succeq 0$, potom $\mathbf{AB} \succeq 0$.

e) Pokud $\mathbf{A} \succ 0$, potom $\mathbf{A}^{-1} \succ 0$.

f) Pokud $\mathbf{A} \succeq 0$ a $\mathbf{B} \succeq 0$, potom $\mathbf{ABA} \succeq 0$.

Které z těchto tvrzení platí? Dokažte nebo najděte protipříklady.

Nápověda ke cvičením

6.6. $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$, $\mathbf{b} = [-1 \ 2]$, $c = 0$

6.13.a) Nemá extrém (má sedlo v bodě $(2, \frac{7}{4})$).

6.13.b) Má minimum v bodě $(3, -1)/2$.

6.14. Nápověda: Je matice symetrická?

6.19. Musí.

Kapitola 7

Rozklad podle singulárních čísel (SVD)

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

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

kde matice $\mathbf{S} \in \mathbb{R}^{m \times n}$ je diagonální (viz definice diagonální matice v §2) s nezápornými prvky a matice $\mathbf{U} \in \mathbb{R}^{m \times m}$ a $\mathbf{V} \in \mathbb{R}^{n \times n}$ jsou ortogonální ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$ a $\mathbf{V}^T\mathbf{V} = \mathbf{I}$).

Diagonální prvky matice \mathbf{S} se nazývají **singulární čísla** matice \mathbf{A} a budeme je značit s_1, \dots, s_p , kde $p = \min\{m, n\}$. Je zvykem je sestupně seřadit,

$$s_1 \geq \dots \geq s_p \geq 0,$$

což lze vždy zajistit permutací sloupců \mathbf{U} a \mathbf{V} . Sloupce matice \mathbf{U} příp. \mathbf{V} se nazývají levé příp. pravé **singulární vektory** matice \mathbf{A} . Rozklad (7.1) se nazývá **rozklad podle singulárních čísel** (*Singular Value Decomposition, SVD*).

Necht' $r = \text{rank } \mathbf{S}$ je počet nenulových singulárních čísel. Dle tvrzení 2 Věty 3.4 (použité na výraz (7.1) a na jeho transpozici) je $\text{rank } \mathbf{A} = \text{rank } \mathbf{S}$. Tedy $\text{rank } \mathbf{A} = r$. Pišme (7.1) jako

$$\mathbf{A} = \underbrace{\begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{S}} \underbrace{\begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}}_{\mathbf{V}^T} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^T \quad (7.2)$$

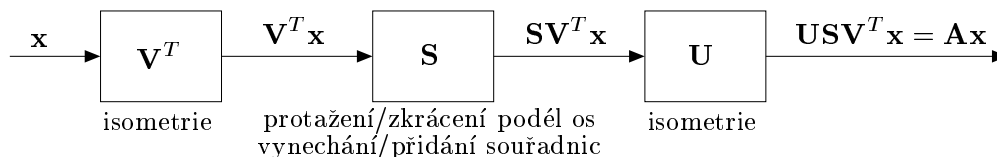
kde $\mathbf{S}_1 = \text{diag}(s_1, \dots, s_r) \in \mathbb{R}^{r \times r}$ je čtvercová diagonální matice, na jejíž diagonále jsou všechna nenulová singulární čísla. Velikosti bloků $\mathbf{U}_1, \mathbf{U}_2, \mathbf{V}_1, \mathbf{V}_2$ a nulových bloků jsou určeny velikostí matice \mathbf{S}_1 (pokud nějaký blok má jeden rozměr nulový, považujeme ho za prázdný). Rozklad $\mathbf{A} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^T$ se nazývá **redukované SVD**.

Redukované SVD se tedy z plného SVD (7.1) získá tak, že matici \mathbf{S} ořízneme na čtvercovou $r \times r$, z matice \mathbf{U} vynecháme posledních $m - r$ sloupců a z matice \mathbf{V} vynecháme posledních $n - r$ sloupců. Plné SVD se z redukovaného SVD získá tak, že úzké matice \mathbf{U}_1 a \mathbf{V}_1 doplníme přidáním sloupců na čtvercové ortogonální, a čtvercovou matici \mathbf{S} doplníme nulami na obdélníkovou stejného rozměru jako \mathbf{A} .

Příklad 7.1. Zde je příklad plného a redukovaného SVD matice 2×3 :

$$\begin{aligned} \mathbf{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & 1/3 \end{bmatrix} = \mathbf{U}\mathbf{S}\mathbf{V}^T \\ &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \end{bmatrix} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^T \quad \square \end{aligned}$$

Vztah (7.1) říká, že každé lineární zobrazení je složením tří jednodušších lineárních zobrazení, a to isometrie \mathbf{V}^T , diagonálního zobrazení \mathbf{S} a isometrie \mathbf{U} . Lineární zobrazení reprezentované diagonální maticí je jednoduše protažení nebo zkrácení podél souřadnicových os, případně (pokud je matice široká) vynechání některých souřadnic nebo (pokud je matice úzká) přidání nulových souřadnic.



V řeči bází to znamená, že pro každé lineární zobrazení lze najít takové ortonormální báze prostoru vzorů a prostoru obrazů, že vzhledem k těmto bázím je zobrazení diagonální.

Motivace SVD: je podobné spektrálního rozkladu, který ne vždy existuje a funguje jen pro čtvercové matice. Platíme tím, že $\mathbf{U} \neq \mathbf{V}$. Dále říci, že SVD je mocný nástroj na zkoumání lineárního zobrazení.

7.1 Existence a výpočet SVD

Dokažme nyní Větu 7.1. Pro jednoduchost uvedeme důkaz jen pro redukované SVD matice s lineárně nezávislými sloupci (tj. \mathbf{A} je čtvercová nebo úzká plně hodnosti).

Důkaz. Protože matice $\mathbf{A}^T \mathbf{A}$ je symetrická, dle spektrálního rozkladu (6.7) existuje diagonální \mathbf{D} a ortogonální \mathbf{V} tak, že

$$\mathbf{A}^T \mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^T.$$

Matice $\mathbf{A}^T \mathbf{A}$ má plnou hodnot (viz tvrzení 2 Věty 3.4) a je pozitivně semidefinitní (viz Cvičení 6.16), tedy \mathbf{D} je regulární s nezápornými prvky. Položme

$$\mathbf{S} = \mathbf{D}^{1/2}, \quad \mathbf{U} = \mathbf{A} \mathbf{V} \mathbf{S}^{-1},$$

kde $\mathbf{D}^{1/2}$ značí diagonální matici splňující $\mathbf{D}^{1/2} \mathbf{D}^{1/2} = \mathbf{D}$ (tj. odmocníme každý diagonální prvek zvlášť). Máme $\mathbf{U}^T \mathbf{U} = \mathbf{S}^{-1} \mathbf{V}^T \mathbf{A}^T \mathbf{A} \mathbf{V} \mathbf{S}^{-1} = \mathbf{S}^{-1} \mathbf{V}^T \mathbf{V} \mathbf{D} \mathbf{V}^T \mathbf{V} \mathbf{S}^{-1} = \mathbf{S}^{-1} \mathbf{D} \mathbf{S}^{-1} = \mathbf{I}$. Platí (7.1), neboť $\mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{A} \mathbf{V} \mathbf{S}^{-1} \mathbf{S} \mathbf{V}^T = \mathbf{A} \mathbf{V} \mathbf{V}^T = \mathbf{A}$. \square

Důkaz ukazuje, jak spočítat SVD ze spektrálního rozkladu matice $\mathbf{A}^T \mathbf{A}$. Tento způsob ale není numericky nejvhodnější, protože výpočet matice $\mathbf{A}^T \mathbf{A}$ může vést ke zbytečným zaokrouhlovacím chybám (viz §5.1.1). Na SVD se proto používají algoritmy, které se explicitnímu výpočtu $\mathbf{A}^T \mathbf{A}$ dokážou vyhnout. Na druhou stranu, pokud nám snížení přesnosti nevádí, počítání SVD spektrálním rozkladem může být rychlejší. Např. pokud je $m \gg n$ a potřebujeme spočítat jen matice \mathbf{V} a \mathbf{S} (a nepotřebujeme \mathbf{U}), spektrální rozklad matice $\mathbf{A}^T \mathbf{A}$ bude typicky rychlejší, protože velikost této matice je malá ($n \times n$).

Matlabský příkaz `[U,S,V]=svd(A)` počítá plné SVD. Redukované SVD není přímo implementované, snadno jej ale dostaneme z příkazu `[U,S,V]=svd(A,'econ')`, který vrátí $\mathbf{U} \in \mathbb{R}^{m \times p}$, $\mathbf{S} \in \mathbb{R}^{p \times p}$ a $\mathbf{V} \in \mathbb{R}^{n \times p}$.

Poznámka o numerické lineární algebře. Potkali jsme již tři rozklady matic: QR, spektrální rozklad, SVD. Je ještě mnoho jiných rozkladů. Návrh numerických algoritmů operace s maticemi, řešení soustav lineárních rovnic a rozklady matic vektorů je předmětem *numerické lineární algebry*. Existují volně dostupné softwarové balíky na numerickou lineární algebru, např. LAPACK a BLAS. Matlab je postaven na balíku LAPACK.

7.2 Ortonormální báze základních podprostorů matice

SVD najde ortonormální báze všech čtyř základních podprostorů generovaných maticí (viz §4.4):

$$\text{rng } \mathbf{U}_1 = \text{rng } \mathbf{A}, \quad (7.3a)$$

$$\text{rng } \mathbf{V}_1 = \text{rng}(\mathbf{A}^T), \quad (7.3b)$$

$$\text{rng } \mathbf{U}_2 = \text{null}(\mathbf{A}^T), \quad (7.3c)$$

$$\text{rng } \mathbf{V}_2 = \text{null } \mathbf{A}. \quad (7.3d)$$

Rovnost (7.3a) platí, neboť $\text{rng } \mathbf{A} = \text{rng}(\mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^T) = \text{rng}(\mathbf{U}_1 \mathbf{S}_1) = \text{rng } \mathbf{U}_1$. To plyne z tvrzení 2 Věty 3.4, neboť \mathbf{V}_1^T a \mathbf{S}_1 mají lineárně nezávislé řádky.

Rovnost (7.3b) plyne z (7.3a), neboť $\mathbf{A}^T = (\mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1)^T = \mathbf{V}_1 \mathbf{S}_1^T \mathbf{U}_1^T = \mathbf{V}_1 \mathbf{S}_1 \mathbf{U}_1^T$.

Protože matice \mathbf{U} a \mathbf{V} jsou ortogonální, z definice ortogonálního doplňku je jasné, že $(\text{rng } \mathbf{U}_1)^\perp = \text{rng } \mathbf{U}_2$ a $(\text{rng } \mathbf{V}_1)^\perp = \text{rng } \mathbf{V}_2$. Rovnosti (7.3c) a (7.3d) nyní plynou z (4.3).

7.3 Nejbližší matice nižší hodnosti

Frobeniova norma matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je číslo

$$\|\mathbf{A}\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{1/2} = \left(\sum_{j=1}^n \|\mathbf{a}_j\|_2^2 \right)^{1/2} \quad (7.4)$$

kde $\mathbf{a}_1, \dots, \mathbf{a}_n$ jsou sloupce matice \mathbf{A} . Protože zjevně $\|\mathbf{A}\|_F = \|\mathbf{A}^T\|_F$, místo sloupců v (7.4) bychom mohli napsat i řádky. Podobně jako eukleidovská norma vektoru, Frobeniova norma se nezmění transformací řádků nebo sloupců matice isometrií, neboli

$$\mathbf{U}^T \mathbf{U} = \mathbf{I}, \quad \mathbf{V}^T \mathbf{V} = \mathbf{I} \quad \implies \quad \|\mathbf{A}\|_F = \|\mathbf{U}\mathbf{A}\|_F = \|\mathbf{A}\mathbf{V}^T\|_F = \|\mathbf{U}\mathbf{A}\mathbf{V}^T\|_F. \quad (7.5)$$

To snadno plyne (promyslete!) z (7.4).

Řešme nyní úlohu, ve které chceme k dané matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ hodnosti r najít nejbližší (ve smyslu Frobeniovy normy) matici \mathbf{A}' dané nižší hodnosti $r' \leq r$. Tedy řešíme úlohu

$$\min \{ \|\mathbf{A} - \mathbf{A}'\|_F \mid \mathbf{A}' \in \mathbb{R}^{m \times n}, \text{rank } \mathbf{A}' = r' \}. \quad (7.6)$$

Řešení je dáno následující větou, kterou uvádíme bez důkazu.

Věta 7.2 (Eckart-Young). Necht' $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ je SVD matice \mathbf{A} . Necht' \mathbf{S}' je diagonální matice, která vznikne vynulováním $r - r'$ nejmenších nenulových diagonálních prvků matice \mathbf{S} (tj. $s'_i = s_i$ pro $i \leq r'$ a $s'_i = 0$ pro $i > r'$). Pak $\mathbf{A}' = \mathbf{U}\mathbf{S}'\mathbf{V}^T$ je řešením úlohy (7.6).

Pokusy o důkaz věty: Dost jasný důkaz je v [Golub], ale pro 2-normu. Dál můžeme zkusit metodu Lagrangeových multiplikátorů. Můžeme použít, že $\|\mathbf{A} - \mathbf{A}'\|_F^2 = \text{tr}((\mathbf{A} - \mathbf{A}')^T (\mathbf{A} - \mathbf{A}')) = \text{tr}(\mathbf{A}^T \mathbf{A}) - 2 \text{tr}(\mathbf{A}^T \mathbf{A}') + \text{tr}(\mathbf{A}'^T \mathbf{A}')$. Celý důkaz spočívá v tom, dokázat, že minimum $\|\mathbf{S} - \mathbf{X}\|_F^2$ za podmínky $\mathbf{X} \in \mathbb{R}^{r \times r}$, $\text{rank } \mathbf{X} = r'$, se nabývá pro diagonální \mathbf{X} (píšu \mathbf{X} místo \mathbf{S}'). Máme

$$\|\mathbf{S} - \mathbf{X}\|_F^2 = \text{tr}(\mathbf{S}^T \mathbf{S}) - 2 \text{tr}(\mathbf{S}^T \mathbf{X}) + \text{tr}(\mathbf{X}^T \mathbf{X}) = \sum_i s_i^2 - 2 \sum_{i,j} s_i x_{ii} + \sum_{i,j} x_{ij}^2.$$

Ekvivalentně můžeme zkusit dokázat, že $\|\mathbf{S}\mathbf{X}\|_F$ za podmínky $\mathbf{X}^T\mathbf{X} = \mathbf{I}$, kde \mathbf{X} je úzká, se minimalizuje pro $\mathbf{X}\mathbf{X}^T$ diagonální. Máme

$$\|\mathbf{S}\mathbf{X}\|_F^2 = \text{tr}(\mathbf{X}^T\mathbf{S}^2\mathbf{X}) = \text{tr}(\mathbf{X}\mathbf{X}^T\mathbf{S}^2) = \text{tr}(\mathbf{P}\mathbf{S}^2) = \sum_i p_i s_i^2$$

kde $\mathbf{P} = \mathbf{X}\mathbf{X}^T$ je ortogonální projektor.

Větu lze formulovat i jinak. Rozklad (7.1) lze napsat jako sumu

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \sum_{i=1}^r s_i \mathbf{u}_i \mathbf{v}_i^T, \quad (7.7)$$

kde $\mathbf{u}_1, \dots, \mathbf{u}_m$ jsou sloupce matice \mathbf{U} a $\mathbf{v}_1, \dots, \mathbf{v}_n$ jsou sloupce matice \mathbf{V} . Všimněte si, že $\mathbf{u}_i \mathbf{v}_i^T \in \mathbb{R}^{m \times n}$ je matice hodnosti 1 (viz §2.5). Matice \mathbf{A}' se získá tak, že z této sumy vezmeme jen prvních r' členů:

$$\mathbf{A}' = \mathbf{U}\mathbf{S}'\mathbf{V}^T = \sum_{i=1}^r s'_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{r'} s_i \mathbf{u}_i \mathbf{v}_i^T.$$

Je-li \mathbf{A}' optimální řešení úlohy (7.6), máme

$$\|\mathbf{A} - \mathbf{A}'\|_F = \|\mathbf{U}\mathbf{S}\mathbf{V}^T - \mathbf{U}\mathbf{S}'\mathbf{V}^T\|_F = \|\mathbf{U}(\mathbf{S} - \mathbf{S}')\mathbf{V}^T\|_F = \|\mathbf{S} - \mathbf{S}'\|_F = (s_{r'+1}^2 + \dots + s_r^2)^{1/2}. \quad (7.8)$$

Vidíme, že ze singulárních čísel lze snadno spočítat, jak je matice daleko od nejbližší matice dané nižší hodnosti.

7.4 Prokládání bodů podprostorem

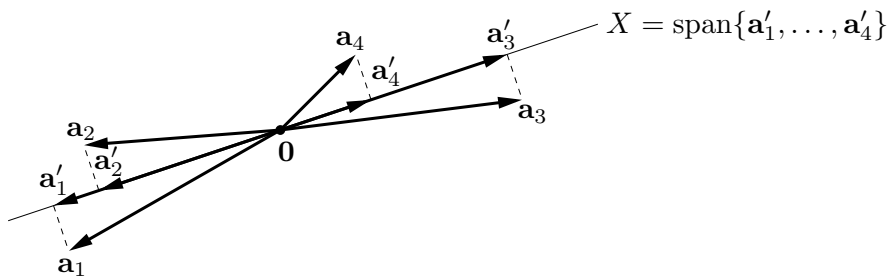
Hledejme (lineární) podprostor $X \subseteq \mathbb{R}^m$ dané dimenze, který minimalizuje součet čtverců vzdáleností k daným bodům¹ $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$. Tuto úlohu nelze nijak převést na úlohu nejmenších čtverců z §5.1. Avšak lze ji řešit pomocí Věty 7.2. Máme

$$\begin{aligned} r &= \text{rank } \mathbf{A} = \dim \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}, \\ r' &= \text{rank } \mathbf{A}' = \dim \text{span}\{\mathbf{a}'_1, \dots, \mathbf{a}'_n\}, \end{aligned}$$

kde \mathbf{a}_j jsou sloupce matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a \mathbf{a}'_j jsou sloupce matice \mathbf{A}' . Dále je

$$\|\mathbf{A} - \mathbf{A}'\|_F^2 = \sum_{j=1}^n \|\mathbf{a}_j - \mathbf{a}'_j\|_2^2.$$

Tedy $X = \text{span}\{\mathbf{a}'_1, \dots, \mathbf{a}'_n\} = \text{rng } \mathbf{A}'$ je takový podprostor dimenze r' , že součet čtverců kolmých vzdáleností bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$ k tomuto podprostoru je minimální:



¹ Ve statistice se této úloze říká *rozvoj podle hlavních komponent* (*principal component analysis, PCA*) nebo *Karhunenův-Loewův rozvoj*.

Obvykle nepotřebujeme najít body \mathbf{a}'_j , ale pouze podprostor X . Jeho ortonormální bázi snadno získáme pomocí vztahů (7.3). Protože pouze prvních r' singulárních čísel matice \mathbf{A}' je nenulových, báze podprostoru $X = \text{rng } \mathbf{A}'$ je množina prvních r' sloupců matice \mathbf{U} v rozkladu $\mathbf{A} = \mathbf{USV}^T$. Někdy může být výhodnější hledat ortogonální doplněk $X^\perp = \text{null}(\mathbf{A}')^\perp$ hledaného podprostoru, jehož báze je posledních $m - r'$ sloupců matice \mathbf{U} .

Příklad 7.2. Máme dáno n bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$ v prostoru \mathbb{R}^3 . Necht' $[\mathbf{a}_1 \cdots \mathbf{a}_n] = \mathbf{USV}^T$ je plné SVD matice, jejíž sloupce jsou tyto body. Označme $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ sloupce matice $\mathbf{U} \in \mathbb{R}^{3 \times 3}$.

Hledejme přímkou procházející počátkem takovou, aby součet čtverců kolmých vzdáleností bodů k této přímce byl co nejmenší. Tato přímka je množina

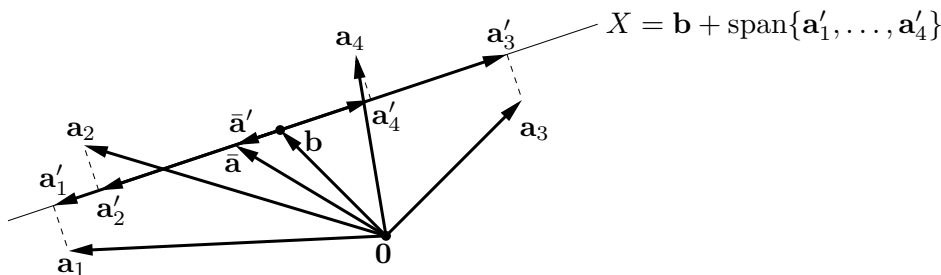
$$\text{span}\{\mathbf{u}_1\} = \{\alpha \mathbf{u}_1 \mid \alpha \in \mathbb{R}\} = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{u}_2^T \mathbf{x} = \mathbf{u}_3^T \mathbf{x} = 0\} = (\text{span}\{\mathbf{u}_2, \mathbf{u}_3\})^\perp.$$

Hledejme rovinu procházející počátkem takovou, aby součet čtverců kolmých vzdáleností bodů k této rovině byl co nejmenší. Tato rovina je množina

$$\text{span}\{\mathbf{u}_1, \mathbf{u}_2\} = \{\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 \mid \alpha_1, \alpha_2 \in \mathbb{R}\} = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{u}_3^T \mathbf{x} = 0\} = (\text{span}\{\mathbf{u}_3\})^\perp. \quad \square$$

7.4.1 Zobecnění na afinní podprostor

Zobecněme nyní úlohu: místo lineárního podprostoru hledáme *afinní* podprostor dimenze r' , který minimalizuje součet čtverců kolmých vzdáleností od bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$. Tento afinní podprostor lze psát jako $X = \mathbf{b} + \text{span}\{\mathbf{a}'_1, \dots, \mathbf{a}'_{r'}\}$ pro nějaké posunutí $\mathbf{b} \in \mathbb{R}^m$ (viz §3.3):



Součet čtverců kolmých vzdáleností k X je (promyslete z obrázku!)

$$\sum_{j=1}^n \|\mathbf{a}_j - \mathbf{a}'_j - \mathbf{b}\|_2^2 = \|\mathbf{A} - \mathbf{A}' - \mathbf{b}\mathbf{1}^T\|_{\mathbb{F}}^2. \quad (7.9)$$

Hledáme $\mathbf{b} \in \mathbb{R}^m$ a $\mathbf{A}' \in \mathbb{R}^{m \times n}$, které minimalizují (7.9) za podmínky $\text{rank } \mathbf{A}' = r'$.

Pokud \mathbf{A}' je pevné, minimalizaci výrazu (7.9) přes proměnnou \mathbf{b} lze snadno vyřešit explicitně (viz Cvičení 5.3): minimum se nabývá v bodě

$$\mathbf{b} = \frac{1}{n} \sum_{j=1}^n (\mathbf{a}_j - \mathbf{a}'_j) = \bar{\mathbf{a}} - \bar{\mathbf{a}}',$$

kde

$$\bar{\mathbf{a}} = \frac{1}{n}(\mathbf{a}_1 + \cdots + \mathbf{a}_n), \quad \bar{\mathbf{a}}' = \frac{1}{n}(\mathbf{a}'_1 + \cdots + \mathbf{a}'_n)$$

značí *těžiště* bodů \mathbf{a}_i resp. \mathbf{a}'_i . Protože $\bar{\mathbf{a}}' \in \text{span}\{\mathbf{a}'_1, \dots, \mathbf{a}'_n\}$, platí $\bar{\mathbf{a}} = \mathbf{b} + \bar{\mathbf{a}}' \in X$. Tedy jsme dokázali, že optimální afinní podprostor X prochází těžištěm $\bar{\mathbf{a}}$ bodů $\mathbf{a}_1, \dots, \mathbf{a}_n$.

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

7.5 Přibližné řešení homogenní soustavy

Řešme homogenní lineární soustavu

$$\mathbf{Ax} = \mathbf{0} \quad (7.10)$$

pro $\mathbf{A} \in \mathbb{R}^{m \times n}$. Množinou řešení je množina $\text{null } \mathbf{A}$, což je lineární podprostor \mathbb{R}^n dimenze $d = n - \text{rank } \mathbf{A}$ (viz (3.12)). Jedno z řešení je vždy $\mathbf{x} = \mathbf{0}$ (tzv. triviální řešení).

Může být homogenní soustava ‘přeuročená’? Přeuročenost soustavy můžeme definovat tak, že dimenze d prostoru řešení je nižší, než nějaká předem daná dimenze $d' > d$. Speciální případ je, když má soustava pouze triviální řešení ($d = 0$), ale my chceme řešení netriviální. Řešme soustavu přibližně tak, že co nejméně změním matici \mathbf{A} , aby prostor řešení měl kýženou dimenzi d' . Neboli nejprve najdeme matici \mathbf{A}' s hodnotí $n - d'$ nejbližší matici \mathbf{A} (dle Věty 7.2) a potom vyřešíme soustavu $\mathbf{A}'\mathbf{x} = \mathbf{0}$.

Vztah k nehomogennímu případu. V §5.1 jsme formulovali přibližné řešení nehomogenní ($\mathbf{b} \neq \mathbf{0}$) soustavy $\mathbf{Ax} = \mathbf{b}$ jako úlohu $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2$. 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 $\mathbf{Ax} = \mathbf{b}$ nemá řešení, změňme vektor \mathbf{b} co nejméně tak, aby soustava řešení měla. Přesněji, hledáme vektor \mathbf{b}' tak, aby pro nějaké \mathbf{x} platilo $\mathbf{Ax} = \mathbf{b}'$ a přitom číslo $\|\mathbf{b} - \mathbf{b}'\|_2$ bylo co nejmenší. Tuto úlohu lze napsat jako

$$\min\{ \|\mathbf{b} - \mathbf{b}'\|_2 \mid \mathbf{Ax} = \mathbf{b}', \mathbf{x} \in \mathbb{R}^n, \mathbf{b}' \in \mathbb{R}^m \}.$$

Zde minimalizujeme přes proměnné \mathbf{x} a \mathbf{b}' (nevadí, že \mathbf{x} se nevyskytuje v účelové funkci). Tato úloha jde zjednodušit (rozmyslete!): dosadíme $\mathbf{b}' = \mathbf{Ax}$ do účelové funkce $\|\mathbf{b} - \mathbf{b}'\|_2$, čímž dostaneme $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2$. Shrňme:

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

7.6 (★) Pseudoinverze obecné matice

Vraťme se k řešení nehomogenní lineární soustavy $\mathbf{Ax} = \mathbf{b}$ pro $\mathbf{A} \in \mathbb{R}^{m \times n}$. V §5 jsme odděleně diskutovali případy, kdy soustava má žádné, jedno nebo nekonečně mnoho řešení. Teď všechny tyto případy spojíme do jediné obecné formulace

$$\min \left\{ \|\mathbf{x}\|_2^2 \mid \mathbf{x} \in \underset{\mathbf{x}' \in \mathbb{R}^n}{\text{argmin}} \|\mathbf{Ax}' - \mathbf{b}\|_2 \right\}. \quad (7.11)$$

To znamená, že hledáme vektor \mathbf{x} , pro který je norma $\|\mathbf{Ax} - \mathbf{b}\|_2$ vektoru residuí minimální; pokud je však takových vektorů více, vybereme z nich takový, který má nejmenší normu $\|\mathbf{x}\|_2$.

Věta 7.3. Necht' $\mathbf{A} = \mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^T$ je redukované SVD matice \mathbf{A} . Pak řešením úlohy (7.11) je $\mathbf{x} = \mathbf{A}^+ \mathbf{b}$, kde

$$\mathbf{A}^+ = \mathbf{V}_1 \mathbf{S}_1^{-1} \mathbf{U}_1^T. \quad (7.12)$$

Důkaz (*). Necht' SVD matice \mathbf{A} je dáno vzorcem (7.2). Platí:

$$\begin{aligned}
\|\mathbf{Ax} - \mathbf{b}\|_2^2 &= \|\mathbf{USV}^T \mathbf{x} - \mathbf{b}\|_2^2 \\
&= \|\mathbf{U}^T(\mathbf{USV}^T \mathbf{x} - \mathbf{b})\|_2^2 && \text{nebot' } \|\mathbf{U}^T \mathbf{z}\|_2 = \|\mathbf{z}\|_2 \text{ pro každé } \mathbf{z} \\
&= \|\mathbf{SV}^T \mathbf{x} - \mathbf{U}^T \mathbf{b}\|_2^2 && \text{nebot' } \mathbf{U}^T \mathbf{U} = \mathbf{I} \\
&= \|\mathbf{S}\mathbf{y} - \mathbf{c}\|_2^2 \\
&= \left\| \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} - \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} \right\|_2^2 \\
&= \left\| \begin{bmatrix} \mathbf{S}_1 \mathbf{y}_1 - \mathbf{c}_1 \\ -\mathbf{c}_2 \end{bmatrix} \right\|_2^2 \\
&= \|\mathbf{S}_1 \mathbf{y}_1 - \mathbf{c}_1\|_2^2 + \|\mathbf{c}_2\|_2^2,
\end{aligned} \tag{7.13}$$

kde jsme označili

$$\mathbf{V}^T \mathbf{x} = \begin{bmatrix} \mathbf{V}_1^T \mathbf{x} \\ \mathbf{V}_2^T \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \mathbf{y}, \quad \mathbf{U}^T \mathbf{b} = \begin{bmatrix} \mathbf{U}_1^T \mathbf{b} \\ \mathbf{U}_2^T \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \mathbf{c}. \tag{7.14}$$

Čeho jsme dosáhli? Ukázali jsme, že výraz $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ je roven výrazu (7.13). Ale ten se mnohem snadněji minimalizuje, protože matice \mathbf{S}_1 je diagonální a regulární. Minimum výrazu (7.13) tedy nastane pro $\mathbf{y}_1 = \mathbf{S}_1^{-1} \mathbf{c}_1$, protože pak bude $\mathbf{S}_1 \mathbf{y}_1 = \mathbf{c}_1$. Protože \mathbf{S}_1 je čtvercová diagonální, její inverze je jednoduše $\mathbf{S}_1^{-1} = \text{diag}(s_1^{-1}, \dots, s_r^{-1})$.

Výraz (7.13) nezávisí na vektoru \mathbf{y}_2 , který tedy můžeme zvolit libovolně. Zvolme jej tak, aby vektor \mathbf{y} měl nejmenší normu. To očividně nastane pro $\mathbf{y}_2 = \mathbf{0}$. Protože $\|\mathbf{y}\|_2 = \|\mathbf{V}^T \mathbf{x}\|_2 = \|\mathbf{x}\|_2$ (plyne z ortogonality \mathbf{V}), bude mít také \mathbf{x} nejmenší normu.

Řešení úlohy (7.11) získáme zpětným dosazením z (7.14):

$$\mathbf{x} = \mathbf{V}\mathbf{y} = [\mathbf{V}_1 \quad \mathbf{V}_2] \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = [\mathbf{V}_1 \quad \mathbf{V}_2] \begin{bmatrix} \mathbf{S}_1^{-1} \mathbf{c}_1 \\ \mathbf{0} \end{bmatrix} = \mathbf{V}_1 \mathbf{S}_1^{-1} \mathbf{c}_1 = \mathbf{V}_1 \mathbf{S}_1^{-1} \mathbf{U}_1^T \mathbf{b} = \mathbf{A}^+ \mathbf{b}. \quad \square$$

Všimněte si, že zatímco v důkazu jsme potřebovali plné SVD, ve vzorci (7.12) vystupuje pouze redukované SVD. Matice \mathbf{U}_2 a \mathbf{V}_2 byly třeba jen pro důkaz.

Matice (7.12) se nazývá **pseudoinverze obecné matice**, přesněji *Moore-Penroseova pseudoinverze*. Pokud \mathbf{A} je čtvercová regulární, (7.12) přejde v obyčejnou inverzi \mathbf{A}^{-1} . Pokud \mathbf{A} je obdélníková s plnou hodnotí, (7.12) souhlasí se vzorci (5.4), (5.10) a (5.11). Ověřte tato tvrzení jako cvičení!

Stopa splňuje $\text{tr } \mathbf{A} = \text{tr}(\mathbf{A}^T)$, $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr } \mathbf{A} + \text{tr } \mathbf{B}$. Dále $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, z čehož plyne $\text{tr}(\mathbf{B}^{-1} \mathbf{AB}) = \text{tr } \mathbf{A}$. Dál $\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}^T \mathbf{A})}$. Viz wiki o stopě.

Nejbližší ortogonální matice, $\min \|\mathbf{A} - \mathbf{R}\|_F$ za podmínky $\mathbf{R}^T \mathbf{R} = \mathbf{I}$. Máme

$$\|\mathbf{A} - \mathbf{R}\|_F^2 = \text{tr}((\mathbf{A} - \mathbf{R})^T (\mathbf{A} - \mathbf{R})) = \text{tr}(\mathbf{A}^T \mathbf{A}) - \text{tr}(\mathbf{A}^T \mathbf{R}) - \text{tr}(\mathbf{R}^T \mathbf{A}) + \text{tr } \mathbf{I}.$$

Ale $\text{tr}(\mathbf{A}^T \mathbf{R}) = \text{tr}(\mathbf{R} \mathbf{A}^T) = \text{tr}(\mathbf{R}^T \mathbf{A}) = \text{tr}(\mathbf{A} \mathbf{R}^T)$. Tedy minimalizace $\|\mathbf{A} - \mathbf{R}\|_F$ je stejná jako maximalizace $\text{tr}(\mathbf{R}^T \mathbf{A})$. Techreport [A Flexible New Technique for Camera Calibration (Appendix C)] ukazuje, jak se to převede na SVD. Máme $\mathbf{A} = \mathbf{USV}^T$. Definujeme ortogonální matici $\mathbf{Z} = \mathbf{V}^T \mathbf{R}^T \mathbf{U}$. Pak

$$\text{tr}(\mathbf{R}^T \mathbf{A}) = \text{tr}(\mathbf{R}^T \mathbf{USV}^T) = \text{tr}(\mathbf{V}^T \mathbf{R}^T \mathbf{US}) = \text{tr}(\mathbf{ZS}) = \sum_i z_{ii} s_i \leq \sum_i s_i.$$

Tedy maximum se dosáhne pro $\mathbf{Z} = \mathbf{I}$, tedy $\mathbf{R} = \mathbf{U}\mathbf{V}^T$.

Ortogonalní Prokrustův problém: $\min \|\mathbf{R}\mathbf{A} - \mathbf{B}\|_F$ za podmínky $\mathbf{R}^T\mathbf{R} = \mathbf{I}$, kde \mathbf{A}, \mathbf{B} jsou široké. Máme

$$\|\mathbf{R}\mathbf{A} - \mathbf{B}\|_F^2 = \text{tr}((\mathbf{R}\mathbf{A} - \mathbf{B})^T(\mathbf{R}\mathbf{A} - \mathbf{B})) = \text{tr}(\mathbf{A}^T\mathbf{A}) - \text{tr}(\mathbf{A}^T\mathbf{R}^T\mathbf{B}) - \text{tr}(\mathbf{B}^T\mathbf{R}\mathbf{A}) + \text{tr}(\mathbf{B}^T\mathbf{B}).$$

Ale $\text{tr}(\mathbf{A}^T\mathbf{R}^T\mathbf{B}) = \text{tr}(\mathbf{B}^T\mathbf{R}\mathbf{A}) = \text{tr}(\mathbf{R}\mathbf{A}\mathbf{B}^T)$. Tedy maximalizujeme $\text{tr}(\mathbf{R}\mathbf{A}\mathbf{B}^T)$, což je to samé jako nalezení nejbližší ortogonální matice k $\mathbf{B}\mathbf{A}^T$.

7.7 Cvičení

7.1. Jsou dány matice

$$\mathbf{A} = \begin{bmatrix} 0.528 & 0.896 & -0.72 \\ -1.204 & -0.528 & 0.96 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} -0.6 & 0.8 \\ 0.8 & 0.6 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} -0.64 & -0.6 & 0.48 \\ -0.48 & 0.8 & 0.36 \\ 0.6 & 0 & 0.8 \end{bmatrix}.$$

Spočítejte matici \mathbf{B} hodnosti jedna takovou, že $\|\mathbf{A} - \mathbf{B}\|_F$ je minimální (kde $\|\cdot\|_F$ značí Frobeniovu normu). Spočítejte hodnotu $\|\mathbf{A} - \mathbf{B}\|_F$ pro tuto matici \mathbf{B} .

7.2. Jsou dány čtyři body $\mathbf{a}_1 = (3, -3, 4)$, $\mathbf{a}_2 = (-2, -3, -2)$, $\mathbf{a}_3 = (1, 0, -1)$, $\mathbf{a}_4 = (3, 1, 0)$ v \mathbb{R}^3 . Najděte množinu $X \subseteq \mathbb{R}^3$, která minimalizuje součet čtverců kolmých vzdáleností bodů k sobě, kde X je

- přímka procházející počátkem,
- rovina procházející počátkem,
- přímka která může ale nemusí procházet počátkem.

Můžete použít počítač.

7.3. Najděte bázi ortogonálního doplňku podprostoru $\text{span}\{(1, 1, 1, -1), (2, -1, -1, 1), (-1, 2, 2, 1)\}$ pomocí SVD. Můžete použít počítač.

7.4. Vyřešte soustavu ze Cvičení 5.2 přibližně ve smyslu nejmenších čtverců pomocí SVD. Můžete použít počítač.

7.5. (★) Dokažte vlastnosti pseudoinverze ze Cvičení 5.19 pomocí vztahu (7.12) pro libovolné (čtvercové či obdélníkové) matice z libovolnou hodnotí.

Nápověda ke cvičením

7.1. $\|\mathbf{A} - \mathbf{B}\|_F = 0.5$

7.2.a) $X = \text{span}\{(0.6912, -0.2181, 0.6889)\}$ (Uvědomte si ale, že báze vašeho podprostoru může být jiná než zde uvedená a přesto podprostor může být stejný.)

7.2.b) $X = \text{span}\{(0.6912, -0.2181, 0.6889), (-0.3771, -0.9221, 0.0864)\}$

7.2.c) $X = (1.25, -1.25, 0.25) + \text{span}\{(0.6599, 0.0280, 0.7508)\}$

7.3. Použijte (7.3c) a $(\text{rng } \mathbf{U}_1)^\perp = \text{rng } \mathbf{U}_2$. Hledaná báze bude třetí sloupec \mathbf{U} až na skalární násobek, tedy např. $(0, -1, 1, 0)$.

Kapitola 8

Nelineární funkce a zobrazení

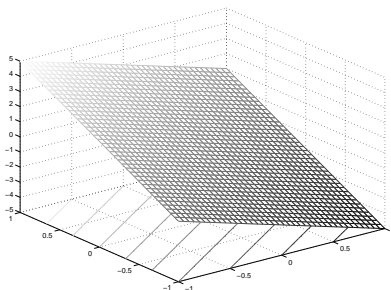
V předchozích kapitolách jsme potkali lineární a afinní zobrazení a kvadratické funkce. V této kapitole si řekneme více o nelineárních funkcích $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a zobrazeních $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (zopakujte si značení funkcí a zobrazení z §1.1.3). Předpokládáme přitom, že student zná analýzu funkce jedné proměnné a úplný základ analýzy funkcí více proměnných.

Dále budeme předpokládat, že definiční obor funkcí a zobrazení je celé \mathbb{R}^n . Tomu tak není vždy, např. definiční obor funkce $f(x) = \sqrt{1-x^2}$ je interval $[-1, 1] \subset \mathbb{R}$. Tento předpoklad ale zjednoduší výklad a pro čtenáře vždy bude snadné látku zobecnit pro jiný definiční obor.

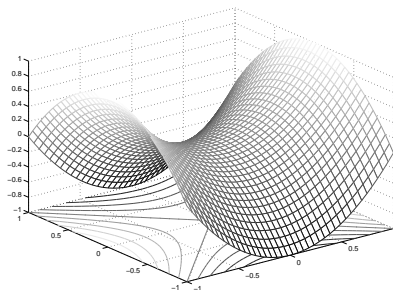
Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ užíváme tyto pojmy:

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

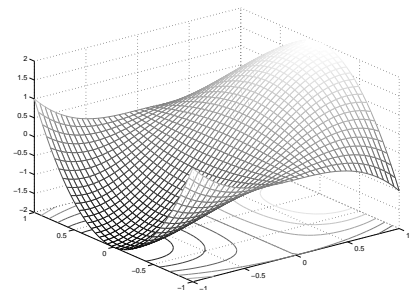
Obrázek ukazuje příklady grafu a vrstevnic funkcí dvou proměnných na obdélníku $[-1, 1]^2$:



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



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



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

Příklad 8.1. Příklady funkcí a zobrazení více proměnných:

1. $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = x^2 - y^2$
2. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = x_1$ (i když x_2, \dots, x_n chybí, f se rozumí jako funkce n proměnných)
3. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = a_1 x_1 + \dots + a_n x_n$ (lineární funkce)
4. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b = a_1 x_1 + \dots + a_n x_n + b$ (afinní funkce)
5. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = e^{-\mathbf{x}^T \mathbf{x}}$ (nenormalizované Gaussovo rozdělení)
6. $f: \mathbb{R}^n \rightarrow \mathbb{R}, f(\mathbf{x}) = \max_{i=1}^n x_i$
7. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^2, \mathbf{f}(t) = (\cos t, \sin t)$ (parametrizace kružnice, množina $\mathbf{f}([0, 2\pi))$ je kružnice)
8. $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^3, \mathbf{f}(t) = (\cos t, \sin t, at)$ (parametrizace šroubovice neboli helixu)

9. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n, \mathbf{f}(\mathbf{x}) = \mathbf{x}$ (identické zobrazení neboli identita)
10. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ (lineární zobrazení)
11. $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ (afinní zobrazení)
12. $\mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \mathbf{f}(u, v) = ((R + r \cos v) \cos u, (R + r \cos v) \sin u, r \sin v)$
(parametrizace toru neboli anuloidu, množina $\mathbf{f}([0, 2\pi) \times [0, 2\pi))$ je torus)
13. Při technice *image morphing* se obrázek např. obličejů zdeformuje na obrázek jiného obličejů. Morphing je realizován zobrazením $\mathbb{R}^2 \rightarrow \mathbb{R}^2$.
14. Elektrické pole přiřadí každému bodu v \mathbb{R}^3 vektor z \mathbb{R}^3 . □

8.1 Spojitost

Neformálně řečeno, zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je **spojité** v bodě $\mathbf{x} \in \mathbb{R}^n$, jestliže bodům blízkým k \mathbf{x} přiřazuje body blízké k $\mathbf{f}(\mathbf{x})$. Abychom tuto větu formalizovali, potřebovali bychom limitu funkce více proměnných, jejíž znalost u čtenáře nepředpokládáme. Uvedeme proto pouze postačující (avšak nikoliv nutnou) podmínku pro spojitost, která nám v praxi postačí. Přitom předpokládáme, že čtenář dokáže ověřit spojitost funkcí jedné proměnné.

Tuto větu na přednáškách nepsat na tabuli (zabere to hodně času), jenom říct.

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' $g: \mathbb{R}^n \rightarrow \mathbb{R}$ je spojitá v bodě \mathbf{x} a $f: \mathbb{R} \rightarrow \mathbb{R}$ je spojitá v bodě $y = g(\mathbf{x})$. Pak složená funkce $f \circ g: \mathbb{R}^n \rightarrow \mathbb{R}$ je spojitá v bodě \mathbf{x} .
4. Necht' funkce $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou spojité v bodě \mathbf{x} . Pak zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ definované jako $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ je spojité v bodě \mathbf{x} .

Příklad 8.2. Z věty snadno ukážeme, že např. funkce dvou proměnných

$$f(x, y) = \sin(x + y^2) \tag{8.1}$$

je spojitá. Podle 1 je x spojitá funkce dvou proměnných (x, y) . Podobně, y^2 je spojitá funkce proměnných (x, y) . Podle 2 je proto funkce $x + y^2$ spojitá. Protože funkce \sin je spojitá, podle 3 je funkce $\sin(x + y^2)$ spojitá. Takto jsme 'rekurzivně' dokázali spojitost celé funkce. □

8.2 Parciální derivace

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

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

kde poslední značení předpokládá, že jsme psali $y = f(\mathbf{x})$. Spočítáme ji tak, že všechny proměnné $x_j, j \neq i$, považujeme za konstanty a zderivujeme funkci podle jediné proměnné x_i .

Příklad 8.3. Parciální derivace funkce (8.1) jsou

$$\frac{\partial f(x, y)}{\partial x} = f_x(x, y) = \cos(x + y^2), \quad (8.2a)$$

$$\frac{\partial f(x, y)}{\partial y} = f_y(x, y) = 2y \cos(x + y^2). \quad (8.2b)$$

□

8.3 Totální derivace

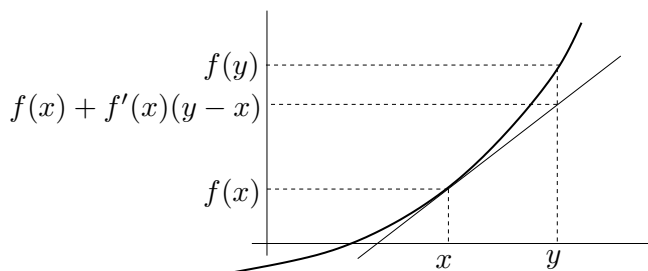
Zopakujme definici derivace funkce jedné proměnné $f: \mathbb{R} \rightarrow \mathbb{R}$ v bodě x . Existuje-li limita

$$\frac{df(x)}{dx} = f'(x) = \lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x}, \quad (8.3)$$

funkce se nazývá *diferencovatelná* v bodě x a hodnota limity se nazývá její *derivace* funkce f v bodě x . Pokud je funkce v bodě x diferencovatelná, lze ji v blízkosti bodu x ‘dobře’ aproximovat afinní funkcí

$$f(y) \approx f(x) + f'(x)(y - x). \quad (8.4)$$

Viz obrázek:



Jak se dá pojem diferencovatelnosti a derivace zobecnit na zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$? Toho nelze dosáhnout zobecněním limity (8.3), ale je lépe vycházet ze vzorce (8.4). Zkusme zobrazení v blízkosti bodu \mathbf{x} aproximovat jako

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

kde symbol $\mathbf{f}'(\mathbf{x})$ označuje matici rozměru $m \times n$, o které ale zatím nevíme nic. Je-li \mathbf{x} pevné, pravá strana výrazu (8.5) je afinní zobrazení v proměnné \mathbf{y} (porovnejte s (3.4): pravou stranu (8.5) lze psát jako $\mathbf{A}\mathbf{y} + \mathbf{b}$, kde $\mathbf{A} = \mathbf{f}'(\mathbf{x})$ a $\mathbf{b} = \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{x}$). Zobrazení je **diferencovatelné** v bodě \mathbf{x} , jestliže je v okolí tohoto bodu ‘podobné’ afinnímu zobrazení, neboli existuje matice $\mathbf{f}'(\mathbf{x})$ taková, že chyba aproximace $\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x})$ je ‘malá’ pro ‘malé’ $\mathbf{y} - \mathbf{x}$. Abychom tuto podmínku formulovali přesně, potřebovali bychom limitu funkce více proměnných, jejíž znalost u čtenáře nepředpokládáme. Ponecháme proto pojem ‘diferencovatelné zobrazení’ nedefinovaný a definujeme pouze o něco silnější vlastnost, která nám v tomto kurzu postačí.

Definice 8.1. Zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je v bodě \mathbf{x} **spojitě diferencovatelné**, jestliže v bodě \mathbf{x} existují všechny parciální derivace $\partial f_i(\mathbf{x})/\partial x_j$ a jsou v tomto bodě spojité.

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

Příklad 8.4. Obě parciální derivace (8.2) funkce (8.1) jsou spojité funkce na celém \mathbb{R}^2 (neboť splňují předpoklady Věty 8.1). Proto je funkce (8.1) spojitě diferencovatelná (a tedy diferencovatelná) na celém \mathbb{R}^2 . \square

Zdůrazněme, že pouhá existence všech parciálních derivací pro diferencovatelnost nestačí.

Příklad 8.5. Necht' je funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definována jako

$$f(x, y) = \begin{cases} 1 & \text{když } x = 0 \text{ nebo } y = 0, \\ 0 & \text{když } x \neq 0 \text{ a } y \neq 0. \end{cases}$$

V bodě $(0, 0)$ existují obě parciální derivace (obě jsou rovny nule). Lze ukázat, že v bodě $(0, 0)$ funkce není diferencovatelná. To nepřekvapí, neboť funkce se v okolí tohoto bodu afinní funkcí vůbec nepodobá. \square

V případě, že je zobrazení \mathbf{f} je v bodě \mathbf{x} diferencovatelné, má matice $\mathbf{f}'(\mathbf{x})$ přirozený tvar: její složky jsou parciální derivace všech složek zobrazení podle všech proměnných:

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

Matice (8.6) se nazývá **totální derivace**¹ (nebo krátce jen **derivace**) zobrazení \mathbf{f} v bodě \mathbf{x} . Z historických důvodů se jí říká také **Jacobiho matice**. Speciální případy:

- Pro $f: \mathbb{R} \rightarrow \mathbb{R}$ je $f'(x)$ *skalár* a splývá s obyčejnou derivací (8.3).
- Pro $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^m$ je $\mathbf{f}'(x)$ *sloupcový vektor*, jehož složky jsou obyčejné derivace složek \mathbf{f} .
- Pro $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je $f'(\mathbf{x})$ *řádkový vektor*.

8.3.1 Derivace složeného zobrazení

Známé 'řetězové pravidlo' pro derivaci složených funkcí jedné proměnné lze přirozeným způsobem rozšířit na zobrazení. Důkaz následující věty není krátký a nebudeme ho uvádět.

Věta 8.3. Necht' $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a $\mathbf{g}: \mathbb{R}^m \rightarrow \mathbb{R}^l$ jsou diferencovatelná zobrazení. Derivace složeného zobrazení $\mathbf{g} \circ \mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^l$ je

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \frac{d\mathbf{g}(\mathbf{f}(\mathbf{x}))}{d\mathbf{x}} = \mathbf{g}'(\mathbf{f}(\mathbf{x})) \mathbf{f}'(\mathbf{x}). \quad (8.7)$$

Dimenze zúčastněných prostorů lze přehledně znázornit diagramem (viz §1.1.2)

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

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

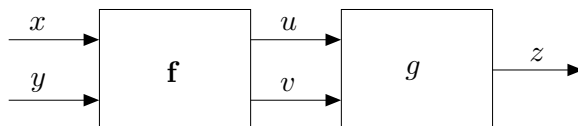
$$\frac{d\mathbf{y}}{d\mathbf{x}} = \frac{d\mathbf{y}}{d\mathbf{u}} \frac{d\mathbf{u}}{d\mathbf{x}}, \quad (8.8)$$

¹Někdy se místo pojmu 'totální derivace' používá pojem 'totální diferenciál'. Tyto pojmy jsou si podobné ale ne identické: totální derivace je *matice* a totální diferenciál je *lineární zobrazení* reprezentované touto maticí.

což se dobře pamatuje, protože du se ‘jakoby vykrátí’ (což ale není důkaz!). Zdůrazněme, že tato rovnost je *násobení matic*. Výraz na levé straně je matice $l \times n$, první výraz na pravé straně je matice $l \times m$ a druhý výraz na pravé straně je matice $m \times n$. Pro $l = m = n = 1$ dostaneme řetězové pravidlo pro derivaci složení funkcí jedné proměnné. Pravidlo se dá zjevným způsobem rozšířit na složení více než dvou zobrazení: *Jacobiho matice složeného zobrazení je součinem Jacobiho matic jednotlivých zobrazení*.

Příklad 8.6. Necht' $g(u, v)$ je diferencovatelná funkce dvou proměnných. Určeme derivaci funkce $z = h(x, y) = g(x + y, xy)$ podle vektoru (x, y) , tedy její parciální derivace podle x a y .

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



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

$$g'(u, v) = \left[\frac{\partial g(u, v)}{\partial u} \quad \frac{\partial g(u, v)}{\partial v} \right] = [g_u(u, v) \quad g_v(u, v)].$$

Derivace zobrazení \mathbf{f} podle vektoru (x, y) je matice 2×2

$$\mathbf{f}'(x, y) = \frac{d\mathbf{f}(x, y)}{d(x, y)} = \begin{bmatrix} \frac{\partial(x+y)}{\partial x} & \frac{\partial(x+y)}{\partial y} \\ \frac{\partial(xy)}{\partial x} & \frac{\partial(xy)}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix}.$$

Derivace zobrazení $h = g \circ \mathbf{f}: \mathbb{R}^2 \rightarrow \mathbb{R}$ podle vektoru (x, y) je matice 1×2 (řádkový vektor)

$$\begin{aligned} \frac{dh(x, y)}{d(x, y)} &= \frac{dg(\mathbf{f}(x, y))}{d(x, y)} = g'(u, v)\mathbf{f}'(x, y) \\ &= [g_u(u, v) \quad g_v(u, v)] \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix} \\ &= [g_u(u, v) + yg_v(u, v) \quad g_u(u, v) + xg_v(u, v)], \end{aligned}$$

kde $u = x + y$ a $v = xy$. □

Příklad 8.7. Ukažme dva způsoby, jak najít parciální derivaci f_x funkce $f(x, y) = e^{(x+y)^2+(xy)^2}$:

- Považujeme y za konstantu a derivujeme f jako funkci jedné proměnné x :

$$f_x = [2(x + y) + 2(xy)y]e^{(x+y)^2+(xy)^2} = 2(x + y + xy^2)e^{(x+y)^2+(xy)^2}.$$

- Položíme $u = x + y$, $v = xy$, $f(u, v) = e^{u^2+v^2}$. Z Příkladu 8.6 máme $f_x = f_u + yf_v$. Jelikož

$$f_u = 2ue^{u^2+v^2}, \quad f_v = 2ve^{u^2+v^2},$$

máme $f_x = f_u + yf_v = 2ue^{u^2+v^2} + y(2v)e^{u^2+v^2} = 2(x + y + xy^2)e^{(x+y)^2+(xy)^2}$. □

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

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

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

□

8.3.2 Derivace maticových výrazů

Jsou-li funkce nebo zobrazení zadány výrazem obsahujícím vektory a matice, derivaci lze vždy spočítat ‘hrubou silou’ tak, že výraz rozepíšeme do složek a spočítáme parciální derivace každé složky podle každé proměnné. Přísně vzato, úkol jsme tím splnili. Je ovšem výhodné tento výsledek zjednodušit tak, že pro něj najdeme maticový výraz.

Příklad 8.9. Odvodme derivaci zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ daného vzorcem $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$. Máme

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

Ale $\partial f_i(\mathbf{x})/\partial x_j = a_{ij}$, tedy dle (8.6) je $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$. Tomu se nedivíme, neboť zobrazení \mathbf{f} je afinní, tedy jeho afinní aproximace (8.5) musí být to samé afinní zobrazení. Oporavdu: pravá strana výrazu (8.5) je rovna $\mathbf{A}\mathbf{y} + \mathbf{b}$ (zkontrolujte!). \square

Příklad 8.10. Odvodme derivaci kvadratické formy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x}$, kde \mathbf{A} je libovolná (ne nutně symetrická) matice velikosti $n \times n$. Napíšeme si funkci f podrobně:

$$\begin{aligned} \mathbf{x}^T \mathbf{A}\mathbf{x} &= a_{11}x_1^2 + a_{21}x_2x_1 + \cdots + a_{n1}x_nx_1 + \\ &\quad a_{12}x_1x_2 + a_{22}x_2^2 + \cdots + a_{n2}x_nx_2 + \\ &\quad \vdots \\ &\quad a_{1n}x_1x_n + a_{2n}x_2x_n + \cdots + a_{nn}x_n^2. \end{aligned}$$

Z tohoto výrazu po troše snahy vidíme, že

$$\frac{\partial f(\mathbf{x})}{\partial x_1} = 2a_{11}x_1 + (a_{21} + a_{12})x_2 + \cdots + (a_{n1} + a_{1n})x_n$$

a podobně pro derivace podle ostatních proměnných. Ale tyto parciální derivace lze uspořádat do řádkového vektoru

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

Příklad 8.11. Počítejme derivaci zobrazení $\mathbf{g}(\mathbf{A}\mathbf{x})$ podle \mathbf{x} , kde $\mathbf{g}: \mathbb{R}^l \rightarrow \mathbb{R}^m$ a $\mathbf{A} \in \mathbb{R}^{l \times n}$ jsou dány. Použijme řetězové pravidlo. Máme $\mathbb{R}^n \xrightarrow{\mathbf{f}} \mathbb{R}^l \xrightarrow{\mathbf{g}} \mathbb{R}^m$, kde $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Je $\mathbf{f}'(\mathbf{x}) = \mathbf{A}$, tedy

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

Následující tabulka uvádí derivace často potkávaných zobrazení. Odvodte je jako cvičení!

$\mathbf{f}(\mathbf{x})$	$\mathbf{f}'(\mathbf{x})$	poznámka
\mathbf{x}	\mathbf{I}	$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$
$\mathbf{A}\mathbf{x}$	\mathbf{A}	$\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$
$\mathbf{x}^T \mathbf{x}$	$2\mathbf{x}^T$	$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$
$\mathbf{x}^T \mathbf{A}\mathbf{x}$	$\mathbf{x}^T (\mathbf{A} + \mathbf{A}^T)$	$\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$
$\mathbf{x}^T \mathbf{a} = \mathbf{a}^T \mathbf{x}$	\mathbf{a}^T	$\mathbf{a} \in \mathbb{R}^n$, $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$
$\ \mathbf{x}\ _2$	$\mathbf{x}^T / \ \mathbf{x}\ _2$	$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$
$\mathbf{g}(\mathbf{A}\mathbf{x})$	$\mathbf{g}'(\mathbf{A}\mathbf{x})\mathbf{A}$	$\mathbf{A} \in \mathbb{R}^{l \times n}$, $\mathbf{g}: \mathbb{R}^l \rightarrow \mathbb{R}^m$, $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$
$\mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x})$	$2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$	$\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$
$\mathbf{g}(\mathbf{x})^T \mathbf{h}(\mathbf{x})$	$\mathbf{g}(\mathbf{x})^T \mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$	$\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}$

8.4 Směrová derivace

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

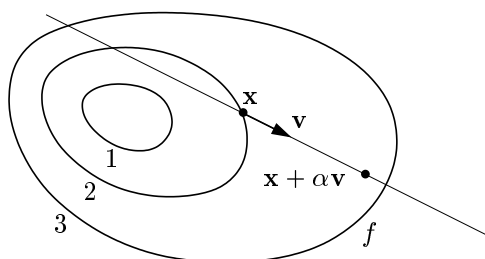
$$\varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha\mathbf{v}). \quad (8.9)$$

Směrová derivace² zobrazení \mathbf{f} v bodě \mathbf{x} ve směru \mathbf{v} je číslo

$$(\varphi'_1(0), \dots, \varphi'_m(0)) = \lim_{\alpha \rightarrow 0} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + \alpha\mathbf{v}) - \mathbf{f}(\mathbf{x})}{\alpha}, \quad (8.10)$$

kde $\varphi'_i(0)$ označuje derivaci i -té složky zobrazení φ v bodě $\alpha = 0$.

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



Směrová derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ve směru i -tého vektoru standardní báze $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ (jednička na i -tém místě) není nic jiného než parciální derivace funkce f podle proměnné x_i .

Věta 8.4. Necht' zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je diferencovatelné v bodě \mathbf{x} . Pak jeho směrová derivace v bodě \mathbf{x} ve směru \mathbf{v} je rovna $\mathbf{f}'(\mathbf{x})\mathbf{v}$.

Důkaz. Zobrazení $\mathbf{y} = \varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha\mathbf{v})$ je složením dvou zobrazení $\mathbf{y} = \mathbf{f}(\mathbf{u})$ a $\mathbf{u} = \mathbf{x} + \alpha\mathbf{v}$. Máme diagram $\mathbb{R} \xrightarrow{\mathbf{u}=\mathbf{x}+\alpha\mathbf{v}} \mathbb{R}^n \xrightarrow{\mathbf{y}=\mathbf{f}(\mathbf{x})} \mathbb{R}^m$. Máme $d\mathbf{u}/d\alpha = \mathbf{v}$. Podle řetězového pravidla je

$$\varphi'(\alpha) = \frac{d\mathbf{y}}{d\alpha} = \frac{d\mathbf{y}}{d\mathbf{u}} \frac{d\mathbf{u}}{d\alpha} = \frac{d\mathbf{f}(\mathbf{u})}{d\mathbf{u}} \mathbf{v}.$$

Pro $\alpha = 0$ je $\mathbf{u} = \mathbf{x}$, čímž je věta dokázána. □

Věta 8.4 říká, že je-li zobrazení \mathbf{f} diferencovatelné, je jeho směrová derivace (v pevném bodě \mathbf{x}) lineární zobrazení směru \mathbf{v} reprezentované Jacobiho maticí $\mathbf{f}'(\mathbf{x})$. Zdůrazněme, že nic takového neplatí, pokud zobrazení \mathbf{f} není diferencovatelné.

Příklad 8.12. Spočítejme směrovou derivaci funkce $f(x, y) = \sin(x + y^2)$ v bodě (x, y) ve směru (u, v) . Podle definice je

$$\begin{aligned} \varphi(\alpha) &= \sin(x + \alpha u + (y + \alpha v)^2), \\ \varphi'(\alpha) &= (u + 2v(y + \alpha v)) \cos(x + \alpha u + (y + \alpha v)^2), \\ \varphi'(0) &= (u + 2vy) \cos(x + y^2). \end{aligned} \quad (8.11)$$

Podle Věty 8.4 je směrová derivace rovna

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

což je stejné jako (8.11). □

Někdy jim říkám, že když směrové derivace v bodě existují ve všech směrech a směrová derivace je lineární funkcí směru, tak zobrazení je diferencovatelné. Ale to není přesně pravda.

² Přesněji jde o *oboustrannou* směrovou derivaci. Jednostrannou směrovou derivaci bychom dostali, kdybychom místo oboustranné limity (8.10) vzali jednostrannou limitu zprava.

8.5 Gradient

Transpozici totální derivace funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ se říká **gradient** a značí se

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

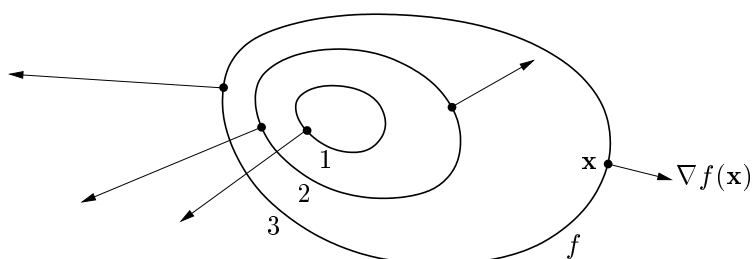
(∇ čteme ‘nabla’). Jelikož $f'(\mathbf{x})$ je řádkový vektor, je gradient sloupcový vektor³.

Zkoumejme směrovou derivaci v pevném bodě \mathbf{x} pro různé normalizované směry \mathbf{v} (tedy $\|\mathbf{v}\|_2 = 1$). Tato derivace je rovna $f'(\mathbf{x})\mathbf{v} = \nabla f(\mathbf{x})^T \mathbf{v}$, což je skalární součin gradientu v bodě \mathbf{x} a vektoru \mathbf{v} . Je jasné (ale promyslete!), že:

- Směrová derivace je největší ve směru $\mathbf{v} = \nabla f(\mathbf{x}) / \|\nabla f(\mathbf{x})\|_2$, tedy když je \mathbf{v} rovnoběžný s gradientem a stejně orientovaný. Tedy směr gradientu je *směr největšího růstu* funkce.
- Velikost gradientu $\|\nabla f(\mathbf{x})\|_2$ je strmost funkce ve směru největšího růstu.
- Směrová derivace ve směru kolmém na gradient je nulová.

Dále lze ukázat (viz diskuze v §9.4.1), že gradient je vždy *kolmý k vrstevnici*.

Obrázek ukazuje tři vrstevnice funkce $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ a její gradienty v několika bodech:



8.6 Parciální derivace druhého řádu

Zderivujeme-li funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ nejdříve podle proměnné x_i a pak podle proměnné x_j , dostaneme parciální derivaci druhého řádu, kterou značíme

$$\frac{\partial}{\partial x_j} \frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}.$$

Je-li $i = j$, píšeme zkráceně

$$\frac{\partial}{\partial x_i} \frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}.$$

Věta 8.5. Pokud jsou druhé parciální derivace

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

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

³ Zavedení nového slova pro transpozici derivace se zdá zbytečné – nicméně důvod je v tom, že totální diferenciál je *lineární funkce*, kdežto gradient je *vektor*. Literatura bohužel není jednotná v rozlišování gradientu a (totální) derivace funkce. Někdy se obojí ztotožňuje a značí $\nabla f(\mathbf{x})$. To ale vede k nekonzistenci se značením používaným v lineární algebře, protože derivace funkce $\mathbb{R}^n \rightarrow \mathbb{R}$ pak není speciálním případem derivace zobrazení $\mathbb{R}^n \rightarrow \mathbb{R}^m$ (tedy Jacobiho matice) pro $m = 1$, což je řádkový vektor.

Příklad 8.13. Určeme všechny druhé parciální derivace funkce $f(x, y) = \sin(x + y^2)$ z Příkladu 8.2. První derivace už tam máme. Nyní druhé derivace:

$$\begin{aligned}\frac{\partial^2 f(x, y)}{\partial x^2} &= \frac{\partial}{\partial x}(\cos(x + y^2)) = -\sin(x + y^2), \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} &= \frac{\partial}{\partial y}(\cos(x + y^2)) = -2y \sin(x + y^2), \\ \frac{\partial^2 f(x, y)}{\partial y \partial x} &= \frac{\partial}{\partial x}(2y \cos(x + y^2)) = -2y \sin(x + y^2), \\ \frac{\partial^2 f(x, y)}{\partial y^2} &= \frac{\partial}{\partial y}(2y \cos(x + y^2)) = -4y^2 \sin(x + y^2).\end{aligned}$$

Vidíme, že vskutku nezáleží na pořadí derivování podle x a podle y . □

Pro funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ budeme značit matici všech druhých parciálních derivací

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

Je to symetrická matice velikosti $n \times n$, která se často nazývá **Hessova matice**.

Co by byla druhá derivace zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$? Nebyla by to již dvojrozměrná tabulka (tedy matice) velikosti $n \times n$, nýbrž třírozměrná tabulka velikosti $m \times n \times n$.

8.7 Taylorův polynom

Necht' funkce $f: \mathbb{R} \rightarrow \mathbb{R}$ má v bodě x derivace až do řádu k . Její **Taylorův polynom** stupně k v bodě x je polynom $T_k: \mathbb{R} \rightarrow \mathbb{R}$ takový, že v bodě x má všechny derivace až do řádu k stejné jako funkce f . V tomto smyslu je polynom T_k aproximací funkce f v okolí bodu x .

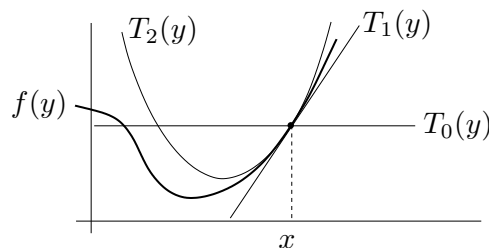
Taylorův polynom je tímto požadavkem definován jednoznačně a má tvar (odvod'te!)

$$T_k(y) = \sum_{i=0}^k \frac{1}{i!} f^{(i)}(x) (y - x)^i, \quad (8.12)$$

kde $f^{(i)}$ označuje i -tou derivaci funkce f (kde nultá derivace je funkce sama, $f^{(0)} = f$) a kde klademe $0! = 1$. Tvary polynomu až do stupně 2:

$$\begin{aligned}T_0(y) &= f(x), \\ T_1(y) &= f(x) + f'(x)(y - x), \\ T_2(y) &= f(x) + f'(x)(y - x) + \frac{1}{2}f''(x)(y - x)^2.\end{aligned}$$

Taylorův polynom nultého stupně T_0 je hodně špatná aproximace, rovná jednoduše konstantní funkci. Polynom prvního stupně $T_1(x)$ už známe ze vzorce (8.4). Polynom druhého stupně T_2 je parabola, která má s funkcí f v bodě x společnou hodnotu a první dvě derivace. Viz obrázek:



Jak zobecnit Taylorův polynom pro funkci více proměnných $f: \mathbb{R}^n \rightarrow \mathbb{R}$? Definujme Taylorův polynom k -tého stupně (funkce f v okolí bodu \mathbf{x}) jako polynom $T_k: \mathbb{R}^n \rightarrow \mathbb{R}$, který má s funkcí f v bodě \mathbf{x} společně všechny parciální derivace až do řádu k . Nebudeme uvádět vzorec pro polynom libovolného stupně, napíšeme jen polynomy do stupně dva:

$$T_0(\mathbf{y}) = f(\mathbf{x}), \quad (8.13a)$$

$$T_1(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.13b)$$

$$T_2(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})(\mathbf{y} - \mathbf{x}). \quad (8.13c)$$

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 (8.13b) je afinní a funkce (8.13c) je kvadratická.

Taylorův polynom lze zobecnit na zobrazení $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tak, že vezmeme Taylorovy polynomy všech složek f_1, \dots, f_m . Polynom prvního stupně tak vede na zobrazení

$$\mathbf{T}_1(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad (8.14)$$

což není nic jiného než (8.5). Polynom druhého stupně vede na zobrazení \mathbf{T}_2 , jehož složky jsou funkce (8.13c). To nejde napsat v maticové formě, protože všech $m \times n \times n$ druhých parciálních derivací se 'nevejde' do matice.

Příklad 8.14. Najděte Taylorův polynom druhého stupně funkce $f(x, y) = \sin(x + y^2)$ z Příkladu 8.3 v bodě $(x_0, y_0) = (2, 1)$.

Máme

$$f(x_0, y_0) = \sin 3,$$

$$f'(x_0, y_0) = \cos(x + y^2) \begin{bmatrix} 1 & 2y \end{bmatrix} \Big|_{(x,y)=(2,1)} = (\cos 3) \begin{bmatrix} 1 & 2 \end{bmatrix},$$

$$f''(x_0, y_0) = -\sin(x + y^2) \begin{bmatrix} 1 & 2y \\ 2y & 4y^2 - 2 \end{bmatrix} \Big|_{(x,y)=(2,1)} = -(\sin 3) \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}.$$

Tedy dle (8.13c) máme (pozor, naše proměnné jsou označené jinak než v (8.13c))

$$\begin{aligned} T_2(x, y) &= f(x_0, y_0) + f'(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T f''(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= \sin 3 + (\cos 3) \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} - \frac{\sin 3}{2} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x - 2 \\ y - 1 \end{bmatrix} \\ &= \sin 3 + (\cos 3)(x - 2 + 2(y - 1)) - \frac{\sin 3}{2}((x - 2)^2 + 4(x - 2)(y - 1) + 2(y - 1)^2) \\ &= (\cos 3)(x + 2y - 4) + (\sin 3)(-x^2/2 - y^2 - 2xy + 3x + 6y + 6). \quad \square \end{aligned}$$

8.8 Cvičení

8.1. Máme množiny $X = [-1, 1] \times \{0\} = \{(x, 0) \mid -1 \leq x \leq 1\} \subseteq \mathbb{R}^2$ a $Y = [-1, 1] \times [-1, 1]$. Načrtněte následující množiny:

a) $\{\mathbf{y} \in \mathbb{R}^2 \mid \min_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|_2 \leq 1\}$

b) $\{\mathbf{y} \in \mathbb{R}^2 \mid \max_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|_2 \leq 2\}$

c) vrstevnice výšky 1 funkce $f(\mathbf{x}) = \min_{\mathbf{x} \in Y} \|\mathbf{x} - \mathbf{y}\|_2$

d) vrstevnice výšky $\sqrt{2}$ funkce $f(\mathbf{x}) = \max_{\mathbf{x} \in Y} \|\mathbf{x} - \mathbf{y}\|_2$

8.2. Je dána funkce dvou proměnných $f(x, y)$.

a) Spočítejte derivace f podle polárních souřadnic (φ, r) , kde $x = r \cos \varphi$, $y = r \sin \varphi$.

b) Bod (x, y) se v čase t pohybuje po křivce dané rovnicí $(x, y) = (t^2 + 2t, \ln(t^2 + 1))$. Najděte derivaci f podle času.

8.3. Spočítejte derivaci funkce $g(\mathbf{u}) = f(\mathbf{a}^T \mathbf{u}, \|\mathbf{u}\|_2)$ podle vektoru \mathbf{u} .

8.4. Nadmořská výška krajiny je dána vzorcem $h(d, s) = 2s^2 + 3sd - d^2 + 5$, kde d je zeměpisná délka (zvětšuje se od západu k východu) a s je zeměpisná šířka (zvětšuje se od jihu k severu). V bodě $(s, d) = (1, -1)$ určete

a) směr nejstrmějšího stoupání terénu

b) strmost terénu v jihovýchodním směru.

8.5. Spočítejte druhou derivaci $f''(x, y)$ (tj. Hessovu matici) funkcí

a) $f(x, y) = e^{-x^2 - y^2}$

b) $f(x, y) = \ln(e^x + e^y)$

8.6. Hessova matice kvadratické formy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je $f''(\mathbf{x}) = \mathbf{A} + \mathbf{A}^T$. Odvod'te.

8.7. Je dána funkce $f(x, y) = 6xy^2 - 2x^3 - 3y^3$. V bodě $(x_0, y_0) = (1, -2)$ najděte Taylorův polynom prvního a druhého stupně.

8.8. *Metoda konečných diferencí* počítá derivaci funkce přibližně jako

$$f'(x) \approx \frac{f(x+h) - f(x)}{h},$$

kde h je malé číslo (dobrá volba je $h = \sqrt{\varepsilon}$, kde ε je strojová přesnost). Toto jde použít i na parciální derivace. Vymyslete si dvě zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a $\mathbf{f}: \mathbb{R}^m \rightarrow \mathbb{R}^l$ pro nějaké navzájem různé dimenze $n, m, l > 1$. Zvolte bod $\mathbf{x} \in \mathbb{R}^n$. Spočítejte přibližně totální derivace (Jacobiho matice) $\mathbf{g}'(\mathbf{x})$ a $\mathbf{f}'(\mathbf{g}(\mathbf{x}))$ v Matlabu metodou konečných diferencí. Potom spočítejte derivaci složeného zobrazení $(\mathbf{f} \circ \mathbf{g})'(\mathbf{x})$ jednak metodou konečných diferencí a jednak vynásobením matic $\mathbf{g}'(\mathbf{x})$ a $\mathbf{f}'(\mathbf{g}(\mathbf{x}))$. Porovnejte.

Kapitola 9

Analytické podmínky na lokální extrémy

9.1 Vlastnosti podmnožin \mathbb{R}^n

Pro $\varepsilon > 0$ a $\mathbf{x} \in \mathbb{R}^n$ se množina

$$U_\varepsilon(\mathbf{x}) = \{ \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| < \varepsilon \} \quad (9.1)$$

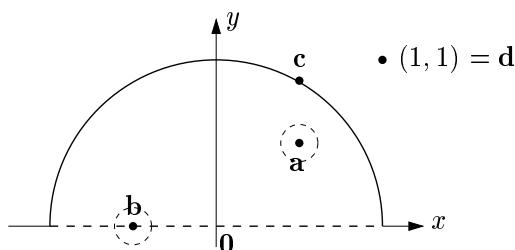
nazývá¹ ε -okolí bodu \mathbf{x} . Je to koule (bez hranice) se středem \mathbf{x} a nenulovým poloměrem ε .

Definice 9.1. Mějme množinu $X \subseteq \mathbb{R}^n$. Bod $\mathbf{x} \in \mathbb{R}^n$ se nazývá její

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

Všimněte si, že hraniční a hromadný bod množiny nemusí patřit do této množiny. Pokud leží bod v množině, je buď vnitřní nebo hraniční, ale ne obojí najednou (dokažte!). **Vnitřek** [hranice] množiny je množina všech jejích vnitřních [hraničních] bodů.

Příklad 9.1. Máme množinu $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, y > 0 \} \cup \{ (1, 1) \} \subseteq \mathbb{R}^2$ na obrázku:



Bod **a** je vnitřní bod množiny, protože existuje $\varepsilon > 0$ takové, že okolí $U_\varepsilon(\mathbf{a})$ celé leží v množině. Bod **b** je hraniční, protože okolí $U_\varepsilon(\mathbf{b})$ má pro každé $\varepsilon > 0$ 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. \square

Příklad 9.2. Bod $1/2$ je vnitřní bod intervalu $(0, 1] \subseteq \mathbb{R}$ a body 0 a 1 jsou hraniční. \square

Příklad 9.3. Množina $[0, 1] \times \{1\} = \{ (x, y) \mid 0 \leq x \leq 1, y = 1 \} \subseteq \mathbb{R}^2$ (úsečka v rovině) nemá žádné vnitřní body. Všechny její body jsou hraniční. Je tedy sama svou vlastní hranicí. \square

¹ Norma v (9.1) může být eukleidovská, ale i libovolná jiná vektorová p -norma (viz §12.5.1). Vnitřek a hranice množiny na výběru normy nezávisí.

9.2 Lokální extrémy

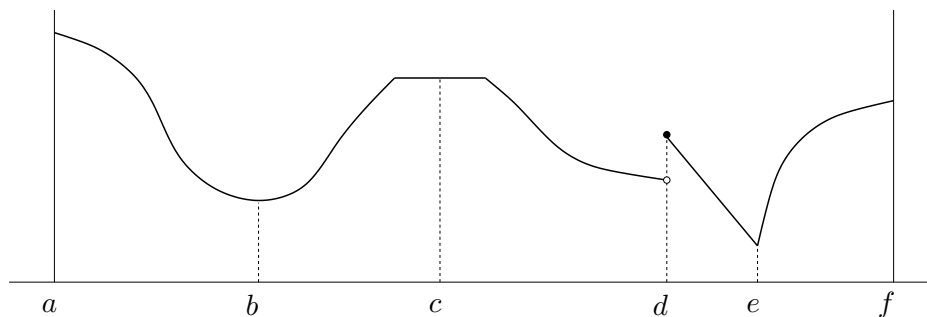
Zopakujte si pojem extrému funkce na množině z §1.1.4!

Definice 9.2. Mějme funkci $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a množinu $X \subseteq \mathbb{R}^n$. Bod $\mathbf{x} \in X$ se nazývá **lokální minimum** funkce f na množině X , existuje-li $\varepsilon > 0$ takové, že \mathbf{x} je minimum funkce f na množině $U_\varepsilon(\mathbf{x}) \cap X$.

Pridat ostre lokální minimum. To je snadné, definovali-li jsme ostře minimum na množině.

Lokální maximum se definuje obdobně. Každé minimum funkce f na množině X je zároveň lokální minimum funkce f na množině X (naopak to ale obecně neplatí). Mluvíme-li o lokálních extrémech, pro zdůraznění někdy ‘obyčejné’ extrémy (ve smyslu §1.1.4) nazýváme **globální extrémy**. Pokud odkaz na množinu X chybí, myslí se jí celý definiční obor funkce f .

Příklad 9.4. Funkce jedné proměnné na obrázku má na uzavřeném intervalu $[a, f]$ v bodě a lokální a zároveň globální maximum, v bodě b lokální minimum, v bodě c lokální maximum a zároveň lokální minimum, v bodě d lokální maximum, v bodě e lokální a zároveň globální minimum, v bodě f lokální maximum. □



Příklad 9.5. Funkce $f(x, y) = x + y$ má na množině $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ globální (a tedy i lokální) minimum v bodě $(1, 1)/\sqrt{2}$ (nakreslete si obrázek!). □

Příklad 9.6. 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. □

9.3 Volné lokální extrémy

Věta 9.1. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{x} \in X \subseteq \mathbb{R}^n$. Necht'

- funkce f je v bodě \mathbf{x} diferencovatelná,
- \mathbf{x} je vnitřní bod množiny X ,
- \mathbf{x} je lokální extrém funkce f na množině X .

Pak $f'(\mathbf{x}) = \mathbf{0}$, neboli všechny parciální derivace funkce f v bodě \mathbf{x} jsou nulové.

Důkaz. Z Definice 9.2 plyne, že existuje $\varepsilon > 0$ tak, že funkce f má v bodě \mathbf{x} (globální) extrém na okolí $U_\varepsilon(\mathbf{x})$. Z toho ovšem plyne, že řez $\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{v})$ funkce f (viz §8.4) v libovolném směru $\mathbf{v} \neq \mathbf{0}$ má (globální) extrém v bodě $\alpha = 0$ na množině $\{\alpha \in \mathbb{R} \mid |\alpha| \leq \varepsilon/\|\mathbf{v}\|\}$. Tedy funkce φ má v bodě $\alpha = 0$ lokální extrém. Tedy její derivace je v tomto bodě nulová (to víme z analýzy funkcí jedné proměnné). Ale tato derivace je směrová derivace funkce f v bodě \mathbf{x} ve směru \mathbf{v} . Parciální derivace jsou speciálním případem směrové derivace. □

Bod, ve kterém má funkce všechny parciální derivace nulové, se nazývá její **stacionární bod**. Věta 9.1 svádí k tomu, aby se použila v situacích, kdy nejsou splněny její předpoklady. Uved'me příklady tohoto chybého použití.

Příklad 9.7. V Příkladu 9.4 jsou předpoklady Věty 9.1 splněny pouze pro body b, c , které jsou stacionární a vnitřní. Body a, f jsou hraniční (tedy ne vnitřní) body intervalu $[a, f]$ a v bodech d, e není funkce diferencovatelná. \square

Příklad 9.8. Funkce $f(x) = x^3$ má na \mathbb{R} v bodě 0 stacionární bod, ale nemá tam lokální extrém. To není v rozporu s Větou 9.1. \square

Příklad 9.9. Funkce $f(\mathbf{x}) = \|\mathbf{x}\|_2$ má na hyperkrychli $X = \{\mathbf{x} \in \mathbb{R}^n \mid -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$ v bodě $\mathbf{0}$ volné lokální minimum (nakreslete si množinu a vrstevnice funkce pro $n = 1$ a pro $n = 2!$). Nemá tam ale stacionární bod, protože tam není diferencovatelná. Dále má funkce na množině X lokální maxima ve všech rozích hyperkrychle, např. v bodě $\mathbf{1}$. V bodě $\mathbf{1}$ ale není stacionární bod, což není v rozporu s Větou 9.1, protože $\mathbf{1}$ není vnitřní bod X . \square

Věta 9.1 říká, že stacionární body jsou body 'podezřelé' z volného lokálního extrému. Udává podmínku *prvního řádu* na volné extrémy, protože obsahuje první derivace. Následující podmínka *druhého řádu* pomůže zjistit, zda je stacionární bod lokálním extrémem, případně jakým.

Věta 9.2. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{x} \in X \subseteq \mathbb{R}^n$. Necht'

- funkce f je v bodě \mathbf{x} dvakrát diferencovatelná,
- \mathbf{x} je vnitřní bod množiny X ,
- $f'(\mathbf{x}) = \mathbf{0}$.

Pak platí:

- Je-li Hessova matice $f''(\mathbf{x})$ pozitivně [negativně] definitní, pak \mathbf{x} je lokální minimum [maximum] funkce f na množině X .
- Je-li $f''(\mathbf{x})$ indefinitní, pak \mathbf{x} není lokální extrém funkce f na množině X .

I když Větu 9.2 nebudeme dokazovat, základní myšlenka důkazu není překvapující. Místo funkce f vyšetřujeme v blízkosti bodu \mathbf{x} její Taylorův polynom druhého stupně (8.13c),

$$T_2(\mathbf{y}) = f(\mathbf{x}) + \underbrace{f'(\mathbf{x})(\mathbf{y} - \mathbf{x})}_{\mathbf{0}} + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

Protože $f'(\mathbf{x}) = \mathbf{0}$, lineární člen je nulový a polynom je tedy kvadratická forma posunutá do bodu \mathbf{x} . Rozdíl je ale v tom, že pokud je kvadratická forma (pozitivně či negativně) semidefinitní, má v počátku extrém, zatímco Věta 9.2 o případě, kdy je $f''(\mathbf{x})$ semidefinitní, nic nepraví. V tom případě v bodě \mathbf{x} lokální extrém být může nebo nemusí (příkladem jsou funkce $f(x) = x^3$ a $f(x) = x^4$ v bodě $x = 0$). Bod \mathbf{x} , ve kterém je $f'(\mathbf{x}) = \mathbf{0}$ a matice $f''(\mathbf{x})$ je indefinitní, se nazývá **sedlový bod**.

Příklad 9.10. Extrémy kvadratické funkce (6.10) umíme hledat pomocí rozkladu na čtverec. Ovšem je to také možné pomocí derivací. Podmínka stacionarity je

$$\frac{d}{d\mathbf{x}}(\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c) = 2\mathbf{x}^T \mathbf{A}^T + \mathbf{b}^T = \mathbf{0}.$$

Po transpozici dostaneme rovnici (6.12a). Druh extrému určíme podle druhé derivace (Hessiánu), který je roven $2\mathbf{A}$ (předpokládáme symetrii \mathbf{A}). To souhlasí s klasifikací extrémů kvadratické formy z §6. \square

9.4 Lokální extrémy vázané rovnostmi

Hledejme minimum funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \}, \quad (9.2)$$

kde $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$. To odpovídá úloze (1.4) s omezeními typu rovnosti:

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmíněk} \quad & g_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, m. \end{aligned} \quad (9.3)$$

Mluvíme o minimu funkce f *vázaném rovnostmi* $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Množina X obsahuje všechna řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, což je soustava m rovnic o n neznámých. Množina X obvykle nemá žádné vnitřní body, proto nelze použít Větu 9.1. V některých případech ale lze vyjádřit všechna řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ parametricky a úlohu tak převést na úlohu bez omezení. Přesněji, existuje množina $Z \subseteq \mathbb{R}^\ell$, která má vnitřní body, a zobrazení $\varphi: \mathbb{R}^\ell \rightarrow \mathbb{R}^n$ tak, že $X = \varphi(Z)$. Pak úlohu převedeme na úlohu $\min\{f(\varphi(\mathbf{z})) \mid \mathbf{z} \in Z\}$, na což je možné užít Větu 9.1. Toto jsme použili v Příkladu 1.2, uveďme další příklady.

Příklad 9.11. Hledejme obdélník s jednotkovým obsahem a minimálním obvodem. Tedy minimalizujeme funkci $f(x, y) = x + y$ za podmínky $xy = 1$, neboli hledáme minima f na množině $X = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 1 - xy = 0\}$.

Množina X nemá žádné vnitřní body (dokažte!), proto nelze použít Větu 9.1. Z podmínky ale máme $y = 1/x$, což dosazeno do účelové funkce dá $f(x, 1/x) = x + 1/x$. Dle Věty 9.1 má tato funkce na svém definičním oboru dva stacionární body $x = \pm 1$. Tedy body podezřelé z lokálního extrému jsou $(x, y) = \pm(1, 1)$. \square

Příklad 9.12. Řešme úlohu

$$\begin{aligned} \min \quad & x + y \\ \text{za podmínky} \quad & x^2 + y^2 = 1, \end{aligned} \quad (9.4)$$

tedy hledáme minimum funkce $f(x, y) = x + y$ na množině $X = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$ kde $g(x, y) = 1 - x^2 - y^2$.

Množina X nemá žádné vnitřní body. Ovšem lze ji parametrizovat jako $X = \{(\cos z, \sin z) \mid z \in [-\pi, \pi]\}$. Úlohu tak převedeme na hledání lokálních extrémů funkce jedné proměnné $f(\cos z, \sin z) = \cos z + \sin z$. Podmínka stacionarity $df(\cos z, \sin z)/dz = -\sin z + \cos z = 0$ má dvě řešení $z = \pm \frac{\pi}{2}$. Tedy body podezřelé z lokálního extrému jsou $(x, y) = \pm \frac{\sqrt{2}}{2}(1, 1)$. \square

Někdy ovšem množinu (9.2) parametrizovat nejde nebo je to složité. Nyní proto odvodíme obecnější postup, *metodu Lagrangeových multiplikátorů*.

9.4.1 Podmínky prvního řádu

Dále budeme předpokládat, že zobrazení \mathbf{g} je spojitě diferencovatelné. V tom případě je množina X ‘zakřivený hladký povrch’² v \mathbb{R}^n . Je-li množina X v okolí nějakého bodu $\mathbf{x} \in X$ ‘hladký povrch’, existuje **tečný prostor** k množině X v bodě \mathbf{x} , což je množina všech vektorů tečných

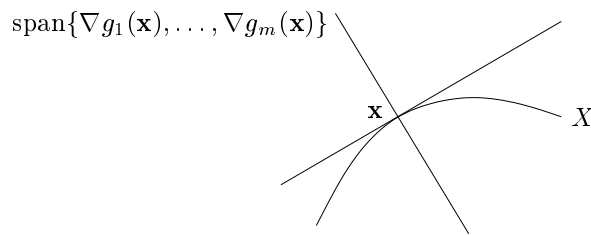
² Přesněji, množina X je příkladem objektu, který se nazývá *diferencovatelný manifold*. Studium takových objektů se zabývá *diferenciální geometrií*.

k povrchu v bodě \mathbf{x} , a **ortogonální prostor** k množině X v bodě \mathbf{x} , což je množina všech vektorů kolmých k povrchu v bodě \mathbf{x} . Tyto dva prostory jsou ortogonální doplněk (viz §4.3) jeden druhého. Zde přesné definice pojmů ‘tečný k povrchu’ a ‘kolmý k povrchu’ neuvádíme a spoléháme na geometrickou intuici.

Lema 9.3. *Necht’ zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je v bodě $\mathbf{x} \in X$ spojitě diferencovatelné. Necht’ Jacobiho matice $\mathbf{g}'(\mathbf{x})$ má hodnost m (neboli gradienty $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$ jsou lineárně nezávislé). Pak ortogonální prostor k množině X v bodě \mathbf{x} je množina*

$$\text{rng } \mathbf{g}'(\mathbf{x})^T = \text{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\}. \quad (9.5)$$

Viz obrázek:



Lemma uvádíme bez důkazu³. Všimněte si, že pro $m = 1$ (tedy pro $g: \mathbb{R}^n \rightarrow \mathbb{R}$) a $\nabla g(\mathbf{x}) \neq \mathbf{0}$ lemma zobecňuje skutečnost, kterou jsme bez důkazu uvedli v §8.5, totiž že gradient funkce je v každém bodě kolmý k její vrstevnici. Lemma ale navíc říká, že *každý* vektor kolmý k vrstevnici musí být násobek gradientu.

Příklad 9.13. Necht’ $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je funkce $g(x, y) = x^2 + y^2 - 1$. Množina X je jednotková kružnice v \mathbb{R}^2 . Máme $\nabla g(x, y) = (2x, 2y)$. Protože pro každé $(x, y) \in X$ je $\nabla g(x, y) \neq (0, 0)$, předpoklady Lematu 9.3 jsou splněny a ortogonální prostor k X v bodě (x, y) je množina $\text{span}\{\nabla g(x, y)\} = \{(\alpha x, \alpha y) \mid \alpha \in \mathbb{R}\}$, což je přímka kolmá ke kružnici. Tečný prostor v bodě (x, y) je ortogonální doplněk této přímky, tedy přímka tečná ke kružnici. \square

Příklad 9.14. Necht’ $g: \mathbb{R}^3 \rightarrow \mathbb{R}$ je funkce $g(x, y, z) = x^2 + y^2 + z^2 - 1$. Množina X je jednotková sféra v \mathbb{R}^3 . Máme $\nabla g(x, y, z) = (2x, 2y, 2z)$. Ortogonální prostor k X v bodě (x, y, z) je množina $\text{span}\{\nabla g(x, y, z)\} = \{(\alpha x, \alpha y, \alpha z) \mid \alpha \in \mathbb{R}\}$, což je přímka kolmá ke sféře. Tečný prostor v bodě (x, y, z) je ortogonální doplněk této přímky, tedy rovina tečná ke sféře. \square

Příklad 9.15. Necht’ $\mathbf{g} = (g_1, g_2): \mathbb{R}^3 \rightarrow \mathbb{R}^2$ je zobrazení

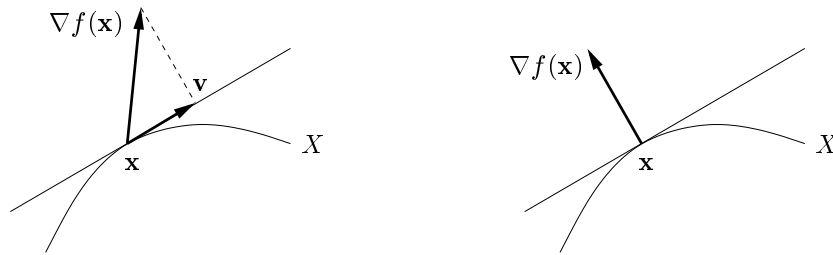
$$\mathbf{g}(x, y, z) = (x^2 + y^2 + z^2 - 1, (x - 1)^2 + y^2 + z^2 - 1).$$

Nulová vrstevnice funkce g_1 je jednotková sféra se středem v bodě $(0, 0, 0)$, nulová vrstevnice funkce g_2 je jednotková sféra se středem v bodě $(1, 0, 0)$. Množina X je průnik těchto dvou sfér, je to tedy kružnice v \mathbb{R}^3 . Máme $\nabla g_1(x, y, z) = 2(x, y, z)$ a $\nabla g_2(x, y, z) = 2(x - 1, y, z)$. Ortogonální prostor k množině X v bodě (x, y, z) je množina $\text{span}\{\nabla g_1(x, y, z), \nabla g_2(x, y, z)\} = \{\alpha_1(x, y, z) + \alpha_2(x - 1, y, z) \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$, což je rovina kolmá ke kružnici v bodě (x, y, z) . Tečný prostor je ortogonální doplněk této množiny, tedy přímka tečná ke kružnici. \square

³ Lema lze dokázat např. pomocí věty o implicitní funkci, která se standardně vyučuje v kursech vícerozměrné analýzy, ale ke které jste se nedostali.

Příklad 9.16. Necht' $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ je funkce $g(x, y) = (x^2 + y^2 - 1)^2$. Množina X je stejná kružnice jako v Příkladě 9.13. Máme $\nabla g(x, y) = 4(x^2 + y^2 - 1)(x, y)$. Pro každý bod $(x, y) \in X$ je $\nabla g(x, y) = (0, 0)$, tedy předpoklady Lematu 9.3 nejsou splněny. Ortogonální prostor není množina $\text{span}\{\nabla g(x, y)\} = \{(0, 0)\}$. \square

Necht' vektor \mathbf{v} označuje průmět gradientu $\nabla f(\mathbf{x})$ do tečného podprostoru k množině X v bodě \mathbf{x} . Směrová derivace funkce f v bodě \mathbf{x} ve směru \mathbf{v} je skalární součin $\nabla f(\mathbf{x})^T \mathbf{v}$. Pokud tato směrová derivace je nenulová (obrázek dole vlevo), bod \mathbf{x} není lokální extrém funkce f na množině X , protože bychom bodem \mathbf{x} mohli infinitezimálně pohnout ve směru vektoru \mathbf{v} a funkci tak zlepšit. Aby byla směrová derivace nulová, musí být $\mathbf{v} = \mathbf{0}$, neboli gradient $\nabla f(\mathbf{x})$ musí být kolmý na tečný podprostor (obrázek vpravo).



Máme tedy tento výsledek: pokud \mathbf{x} je lokální extrém funkce f na množině X , je vektor $\nabla f(\mathbf{x})$ kolmý na tečný podprostor k množině X v bodě \mathbf{x} . Protože tento tečný podprostor je ortogonální doplněk prostoru (9.5), znamená to, že vektor $\nabla f(\mathbf{x})$ patří do množiny (9.5), tedy je lineární kombinací gradientů $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$. Tedy existují čísla $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tak, že

$$\nabla f(\mathbf{x}) + \lambda_1 \nabla g_1(\mathbf{x}) + \dots + \lambda_m \nabla g_m(\mathbf{x}) = \mathbf{0}. \quad (9.6)$$

Výsledek našich úvah se obvykle formuluje následujícím způsobem.

Věta 9.4. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \in X$. Necht'

- f a \mathbf{g} jsou v bodě \mathbf{x} spojitě diferencovatelné,
- matice $\mathbf{g}'(\mathbf{x})$ má hodnost m ,
- bod \mathbf{x} je lokální extrém funkce f na množině X .

Pak existují čísla $(\lambda_1, \dots, \lambda_m) = \boldsymbol{\lambda} \in \mathbb{R}^m$ tak, že $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$, kde funkce $L: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ je dána jako

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \dots + \lambda_m g_m(\mathbf{x}). \quad (9.7)$$

Zápis $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$ označuje, že parciální derivace funkce L podle $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$ jsou nulové, neboli bod $(\mathbf{x}, \boldsymbol{\lambda}) \in \mathbb{R}^{m+n}$ je stacionární bod funkce L . Rovnost $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} = \mathbf{0}$ je ekvivalentní rovnosti (9.6). Rovnost $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \boldsymbol{\lambda} = \mathbf{g}(\mathbf{x}) = \mathbf{0}$ je ekvivalentní omezením. Číslům λ_i se říká **Lagrangeovy multiplikátory** a funkci (9.7) **Lagrangeova funkce**.

Příklad 9.17. Řešme znovu úlohu (9.4). Lagrangeova funkce je $L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2)$. Její stacionární body (x, y, λ) jsou řešeními soustavy tří rovnic o třech neznámých

$$\begin{aligned} \partial L(x, y, \lambda)/\partial x &= 1 - 2\lambda x &= 0 \\ \partial L(x, y, \lambda)/\partial y &= 1 - 2\lambda y &= 0 \\ \partial L(x, y, \lambda)/\partial \lambda &= 1 - x^2 - y^2 &= 0. \end{aligned}$$

První dvě rovnice dají $x = y = 1/(2\lambda)$. Dosazením do třetí máme $2/(2\lambda)^2 = 1$, což dá dva kořeny $\lambda = \pm 1/\sqrt{2}$. Stacionární body funkce L jsou dva, $(x, y, \lambda) = \pm(1, 1, 1)/\sqrt{2}$. Tedy máme dva kandidáty na lokální extrémy, $(x, y) = \pm(1, 1)/\sqrt{2}$.

Tuto jednoduchou úlohu je samozřejmě snadné vyřešit úvahou. Nakreslete si kružnici $X = \{(x, y) \mid x^2 + y^2 = 1\}$ a několik vrstevnic funkce f a najděte kýžené extrémy! \square

Příklad 9.18. Řešme úlohu (9.4), kde ale omezení změňme na $g(x, y) = (1 - x^2 - y^2)^2 = 0$. Podle Příkladu 9.16 máme $g'(x, y) = (0, 0)$ pro každé $(x, y) \in X$, čekáme tedy problém.

Stacionární body Lagrangeovy funkce $L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2)^2$ musí splňovat

$$\begin{aligned}\partial L(x, y, \lambda)/\partial x &= 1 - 4\lambda x(1 - x^2 - y^2) = 0 \\ \partial L(x, y, \lambda)/\partial y &= 1 - 4\lambda y(1 - x^2 - y^2) = 0 \\ \partial L(x, y, \lambda)/\partial \lambda &= (1 - x^2 - y^2)^2 = 0.\end{aligned}$$

Tyto rovnice si odporují. Jelikož $1 - x^2 - y^2 = 0$, tak např. první rovnice říká $1 - 4\lambda x \cdot 0 = 0$, což neplatí pro žádné (x, λ) . Závěr je, že lokální extrémy $(x, y) = \pm(1, 1)/\sqrt{2}$ jsme nenašli. \square

Příklad 9.19. Vraťme se k úloze (5.8), tedy k hledání řešení nehomogenní lineární soustavy s nejmenší normou. Lagrangeova funkce je

$$L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{x}^T \mathbf{x} + 2\boldsymbol{\lambda}^T (\mathbf{b} - \mathbf{A}\mathbf{x}),$$

kde přidaná dvojka nemění situaci. Je $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} = 2\mathbf{x}^T - 2\boldsymbol{\lambda}^T \mathbf{A}$ (odvod'te!). Stacionární body funkce L tedy získáme řešením soustavy (5.9), kterou jsme v 5.2 odvodili úvahou. \square

Předchozí příklad vyžaduje od studenta nejen znalost metody Lagrangeových multiplikátorů, ale i jistou zručnost v manipulaci s maticovými výrazy. Cvičte tuto zručnost ve Cvičnicích 9.22–9.25!

Věta 9.4 udává podmínky prvního řádu na extrémy vázané rovnostmi. Říká, že pokud $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod Lagrangeovy funkce, pak bod \mathbf{x} je 'podezřelý' z lokálního extrému funkce f na množině X . Jak poznáme, zda tento bod je lokální extrém, případně jaký? Podmínky druhého řádu pro vázané extrémy uvádíme nepovinně v §9.4.2. Zde pouze zdůrazníme, že druh lokálního extrému *nelze* zjistit podle definitnosti Hessovy matice $L''(\mathbf{x}, \boldsymbol{\lambda})$, tedy je chybou použít Větu 9.2 na funkci L .

9.4.2 (*) Podmínky druhého řádu

Věta 9.5. Necht' $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$ a $\boldsymbol{\lambda} \in \mathbb{R}^m$. Necht'

- $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod Lagrangeovy funkce, neboli $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} = \mathbf{0}$ a $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \boldsymbol{\lambda} = \mathbf{0}$,
- f a \mathbf{g} jsou dvakrát diferencovatelné v bodě \mathbf{x} .

Pak platí:

- Je-li $\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x}^2$ pozitivně [negativně] definitní na nulovém prostoru matice $\mathbf{g}'(\mathbf{x})$, má f v bodě \mathbf{x} ostré lokální minimum [maximum] vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.
- Je-li $\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x}^2$ indefinitní na nulovém prostoru matice $\mathbf{g}'(\mathbf{x})$, nemá f v bodě \mathbf{x} lokální minimum ani lokální maximum vázané podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

Zde výraz

$$\frac{\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}^2} = f''(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i''(\mathbf{x})$$

značí druhou derivaci (Hessovu matici) funkce $L(\mathbf{x}, \boldsymbol{\lambda})$ podle \mathbf{x} v bodě $(\mathbf{x}, \boldsymbol{\lambda})$. Tvrzení, že matice \mathbf{A} je pozitivně definitní na množině T znamená, že $\mathbf{y}^T \mathbf{A} \mathbf{y} > 0$ pro každé $\mathbf{y} \in T \setminus \{\mathbf{0}\}$.

Jak zjistíme definitnost dané matice \mathbf{A} na nulovém prostoru Jacobiho matice $\mathbf{g}'(\mathbf{x})$? Najdeme-li bázi \mathbf{B} tohoto nulového prostoru, pak každý prvek množiny T lze parametrizovat jako $\mathbf{y} = \mathbf{B} \mathbf{z}$. Protože $\mathbf{y}^T \mathbf{A} \mathbf{y} = \mathbf{z}^T \mathbf{B}^T \mathbf{A} \mathbf{B} \mathbf{z}$, převedli jsme problém na zjišťování definitnosti matice $\mathbf{B}^T \mathbf{A} \mathbf{B}$.

Příklad 9.20. Najdeme strany kvádrů s jednotkovým objem a minimálním povrchem. Tedy minimalizujeme $xy + xz + yz$ za podmínky $xyz = 1$. Lagrangeova funkce je

$$L(x, y, z, \lambda) = xy + xz + yz + \lambda(1 - xyz).$$

Položením derivací L rovným nule máme soustavu

$$\begin{aligned} L'_x(x, y, z, \lambda) &= y + z - \lambda yz = 0 \\ L'_y(x, y, z, \lambda) &= x + z - \lambda xz = 0 \\ L'_z(x, y, z, \lambda) &= x + y - \lambda xy = 0 \\ L'_\lambda(x, y, z, \lambda) &= xyz - 1 = 0. \end{aligned}$$

Soustava je zjevně splněna pro $(x, y, z, \lambda) = (1, 1, 1, 2)$. Máme ukázat, že tento bod odpovídá lokálnímu minimu. Máme

$$\frac{\partial^2 L(x, y, z, \lambda)}{\partial (x, y, z)^2} = \begin{bmatrix} 0 & 1 - \lambda z & 1 - \lambda y \\ 1 - \lambda z & 0 & 1 - \lambda x \\ 1 - \lambda y & 1 - \lambda x & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}. \quad (9.8)$$

Ukážeme, že tato matice je pozitivně definitní na nulovém prostoru Jacobiho matice

$$g'(x, y, z) = [-yz \quad -xz \quad -xy] = [-1 \quad -1 \quad -1].$$

Nejdříve zkusme štěstí, zda matice (9.8) není pozitivně definitní již na \mathbb{R}^3 – v tom případě by zjevně byla pozitivně definitní i na nulovém prostoru $g'(x, y, z)$ (promyslete, proč to tak je!). Není tomu tak, protože její vlastní čísla jsou $\{-2, 1, 1\}$, tedy je indefinitní.

Nějakou bázi nulového prostoru matice $g'(x, y, z)$ snadno najdeme ručně, např.

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Snadno zjistíme, že matice

$$\mathbf{B}^T \frac{\partial^2 L(x, y, z, \lambda)}{\partial (x, y, z)^2} \mathbf{B} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

má vlastní čísla $\{2, 1\}$, tedy je pozitivně definitní. □

9.5 Lokální extrémý vázané nerovnostmi ‘hrubou silou’

Změňme nyní úlohu (9.3) tak, že podmínky budou nerovnosti. Hledáme tedy minimum funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \}, \quad (9.9)$$

kde $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$. To odpovídá úloze

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmíněk} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m. \end{aligned} \quad (9.10)$$

Analogicky k Větě 9.4 existují nutné podmínky prvního pro lokální minimum úlohy (9.10), tzv. *Karush-Kuhn-Tuckerovy (KKT) podmínky*. My je ovšem v tomto kursu nebudeme uvádět. Abychom se ale s omezeními typu nerovnosti trochu seznámili, uvedeme následující naivní algoritmus na řešení úlohy (9.10) za předpokladu, že umíme řešit úlohy s omezeními typu rovnosti.

Uvažujme nejprve případ jediného omezení ($m = 1$), tedy $X = \{ \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \leq 0 \}$. Omezení $g(\mathbf{x}) \leq 0$ nazveme **aktivní** v bodě \mathbf{x} když $g(\mathbf{x}) = 0$, a **neaktivní** když $g(\mathbf{x}) < 0$. Lokální minimum se může nabývat buď ve vnitřním bodě nebo v hraničním bodě množiny X . Pokud je funkce g spojitá, lze ukázat, že vnitřní body množiny X jsou právě ty body, ve kterých je omezení neaktivní, a hraniční body jsou právě ty, ve kterých je omezení aktivní. Pro nalezení kandidátů na lokální extrémý ve vnitřních bodech použijeme Větu 9.1. Pro nalezení kandidátů na lokální extrém v hraničních bodech použijeme Větu 9.4.

Tuto úvahu lze zobecnit pro více omezení, $m \geq 1$. Pokud \mathbf{x} je optimální řešení úlohy (9.10), budou v něm nějaká omezení aktivní a zbylá neaktivní. Pro každou podmnožinu $I \subseteq \{1, \dots, m\}$ najdeme všechna lokální minima funkce f za podmínek $g_i(\mathbf{x}) = 0, i \in I$, a $g_i(\mathbf{x}) < 0, i \notin I$. Musíme tedy umět hledat lokální minima funkce f za omezení typu rovnosti ($g_i(\mathbf{x}) = 0$) a omezení typu ostré nerovnosti ($g_i(\mathbf{x}) < 0$). Pokud jsou funkce g_i spojité, k tomu stačí najít lokální minima funkce f za podmínek typu rovnosti a pak pouze ověřit podmínky typu ostré nerovnosti. **Tohle nutno vylepsit!**

Tento algoritmus má nevýhodu v tom, že musíme vyzkoušet všech 2^m podmnožin I . Proto jej lze použít jen pro velmi malé m .

Příklad 9.21. Hledejme všechny lokální extrémý funkce $f(x, y, z)$ za podmínek

$$\begin{aligned} x^2 + y^2 + z^2 &\leq 1 \\ z &\geq 0 \end{aligned}$$

Máme $g_1(x, y, z) = x^2 + y^2 + z^2 - 1$ a $g_2(x, y, z) = -z$. Množina přípustných řešení je polokoule. Provedeme postupně tyto kroky:

- $I = \emptyset$ (*obě podmínky neaktivní*): Najdeme všechny lokální extrémý funkce f na celém \mathbb{R}^3 . Pro každý nalezený lokální extrém ověříme, zda splňuje podmínky $x^2 + y^2 + z^2 < 1$ a $z > 0$ (tedy leží v půlkouli).
- $I = \{1\}$ (*první podmínka aktivní, druhá neaktivní*): Najdeme všechny lokální extrémý funkce f na sféře $x^2 + y^2 + z^2 = 1$. Pro každý z nich ověříme, zda splňuje podmínku $z > 0$ (tedy leží na správné polovině sféry).
- $I = \{2\}$ (*první podmínka neaktivní, druhá aktivní*): Najdeme všechny lokální extrémý funkce f na rovině $z = 0$. Pro každý z nich ověříme, zda splňuje podmínku $x^2 + y^2 + z^2 < 1$ (tedy leží v kruhu, jenž je průnikem koule a roviny $z = 0$).

- $I = \{1, 2\}$ (obě podmínky aktivní): Najdeme všechny lokální extrémy funkce f za podmínek $x^2 + y^2 = 1$ a $z = 0$ (tedy na kružnici). Nemusíme ověřovat nic. \square

9.6 Cvičení

9.1. Co je vnitřek a hranice těchto množin?

- $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \geq 0\}$
- $\{(x, y) \in \mathbb{R}^2 \mid y = x^2, -1 < x \leq 1\}$
- $\{(x, y) \in \mathbb{R}^2 \mid xy < 1, x > 0, y > 0\}$
- $\{\mathbf{x} \in \mathbb{R}^n \mid \max_{i=1}^n x_i \leq 1\}$
- $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$, kde $\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$ (nadrovina)
- $\{\mathbf{x} \in \mathbb{R}^n \mid b \leq \mathbf{a}^T \mathbf{x} \leq c\}$, kde $\mathbf{a} \in \mathbb{R}^n, b, c \in \mathbb{R}$ (panel)
- $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$, kde \mathbf{A} je široká (afinní podprostor \mathbb{R}^n)

9.2. Je dána funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$, množiny $Y \subseteq X \subseteq \mathbb{R}^n$, a bod $\mathbf{x} \in Y$. Uvažujme dva výroky:

- Funkce f má v bodě \mathbf{x} lokální minimum na množině X .
- Funkce f má v bodě \mathbf{x} lokální minimum na množině Y .

Vyplývá (b) z (a)? Vyplývá (a) z (b)? Dokažte z definice lokálního extrému nebo vyvrát'ete nalezením protipříkladu.

9.3. Může nastat případ, kdy funkce na množině má lokální minimum ale nemá na ní globální minimum? Odpověď dokažte.

9.4. Funkce $f(x, y, z)$ 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

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

9.5. Pro následující funkce spočítejte (na papíře) stacionární body. Pro každý stacionární bod určete, zda je to lokální minimum, lokální maximum, či sedlový bod. Pokud to určit nedokážete, odůvodněte.

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

9.6. Dokažte, že funkce $f(x, y) = x$ nabývá za podmínky $x^3 = y^2$ minima pouze v počátku. Ukažte, že metoda Lagrangeových multiplikátorů toto minimum nenajde.

Následující úlohy se pokuste vyřešit parametrizací podmínek (analogicky k Příkladu 9.12) a pak metodou Lagrangeových multiplikátorů. Pokud jedna z těchto metod není použitelná, vynechte ji. Při použití metody Lagrangeových multiplikátorů stačí pouze najít stacionární body Lagrangeovy funkce – nemusíte určovat, jde-li o lokální extrémy a případně jaké.

9.7. Najděte lokální extrémy funkcí

- 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) $f(x, y) = e^{xy}$

na kružnici $x^2 + y^2 = 1$. Náповěda: Někdy je dobré účelovou funkci zjednodušit, pokud to nezmění řešení.

9.8. Najděte extrémy funkce

- a) $f(x, y, z) = x + yz$ za podmínek $x^2 + y^2 + z^2 = 1$ a $z^2 = x^2 + y^2$
- b) $f(x, y, z) = xyz$ za podmínek $x^2 + y^2 + z^2 = 1$ a $xy + yz + zx = 1$

9.9. Najděte 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) $(\star) f(x, y, z) = x^3 + 2xyz - z^3$

na sféře $x^2 + y^2 + z^2 = 1$.

9.10. Rozložte dané kladné reálné číslo na součin n kladných reálných čísel tak, aby jejich součet byl co nejmenší.

9.11. Spočítejte rozměry tělesa tak, aby mělo při daném objemu nejmenší povrch:

- a) kvádr
- b) kvádr bez víka (má jednu dolní stěnu a čtyři boční, horní stěna chybí)
- c) válec
- d) 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 $S = \pi(R+r)\sqrt{(R-r)^2 + h^2}$. Můžete použít vhodný numerický software na řešení vzniklé soustavy rovnic.

9.12. Najděte bod nejbliže počátku na křivce

- a) $x + y = 1$
- b) $x + 2y = 5$

- c) $y = x^3 + 1$
 d) $x^2 + 2y^2 = 1$

- 9.13. Necht' \mathbf{x}^* je bod nejbližze počátku na nadploše $h(\mathbf{x}) = 0$. Ukažte metodou Lagrangeových multiplikátorů, že vektor \mathbf{x}^* je kolmý k tečné nadrovině plochy v bodě \mathbf{x}^* .
- 9.14. Máme kouli o poloměru r a středu \mathbf{x}_0 , tj. množinu $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\|_2 \leq r\}$. Máme nadrovinu $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$.
- 9.15. Do elipsy o daných délkách os vepište obdélník s maximálním obsahem. Předpokládejte přitom, že strany obdélníku jsou rovnoběžné s osami elipsy.
- 9.16. *Fermatův princip* v paprskové optice říká, že cesta mezi libovolnými dvěma body na paprsku má takový tvar, aby ji světlo proběhlo za čas kratší než jí blízké dráhy. Později se zjistilo, že správným kritériem není *nejkratší* ale *extrémní* čas. Tedy skutečná dráha paprsku musí mít čas větší nebo menší než jí blízké dráhy. Z tohoto principu odvod'te:
- 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

- a) pro rovinné zrcadlo a rovinné rozhraní (což vede na minimalizaci bez omezení),
 b) pro zrcadlo a rozhraní tvaru obecné plochy s rovnicí $g(\mathbf{x}) = 0$. Dokážete najít situaci, kdy skutečná dráha paprsku má čas *větší* než jí blízké dráhy?
- 9.17. Rozdělení pravděpodobnosti diskretní náhodné proměnné je funkce $p: \{1, \dots, n\} \rightarrow \mathbb{R}_+$ (tj. soubor nezáporných čísel $p(1), \dots, p(n)$) splňující $\sum_{x=1}^n p(x) = 1$.
- a) *Entropie* náhodné proměnné s rozdělením p je rovna $-\sum_{x=1}^n p(x) \log p(x)$, kde \log je přirozený logaritmus. Najděte rozdělení s maximální entropií.
 b) Dokážete *Gibbsovu nerovnost* (též zvanou *informační nerovnost*): pro každé dvě rozdělení p, q platí

$$\sum_{x=1}^n p(x) \log q(x) \geq \sum_{x=1}^n p(x) \log p(x),$$

přičemž rovnost nastává jen tehdy, když $p = q$.

- 9.18. (★) Máme trojúhelník se stranami délek a, b, c . Uvažujme bod, který má takovou polohu, že součet čtverců jeho vzdáleností od stran trojúhelníku je nejmenší možný. Jaké budou vzdálenosti x, y, z tohoto bodu od stran trojúhelníku?
- 9.19. (★) Máme krychli s délkou hrany 2. Do stěny krychle je vepsána kružnice (která má tedy poloměr 1) a okolo sousední stěny je opsána kružnice (která má tedy poloměr $\sqrt{2}$). Najděte nejmenší a největší vzdálenost mezi body na kružnicích.
- 9.20. (★) Najděte extrémy funkce

$$f(x, y, z, u, v, w) = (1 + x + u)^{-1} + (1 + y + v)^{-1} + (1 + z + w)^{-1}$$

za podmínek $xyz = a^3$, $uvw = b^3$ a $x, y, z, u, v, w > 0$.

9.21. Popište množinu řešení soustavy

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

Najděte takové řešení soustavy, aby výraz $\sqrt{x^2 + y^2 + z^2}$ byl co nejmenší. Najděte co nejvíce způsobů řešení.

- 9.22. Minimalizujte $\mathbf{x}^T \mathbf{x}$ za podmínky $\mathbf{a}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?
- 9.23. Maximalizujte $\mathbf{a}^T \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{x} = 1$. Jaký je geometrický význam úlohy?
- 9.24. Minimalizujte $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\mathbf{b}^T \mathbf{x} = 1$, kde \mathbf{A} je pozitivně definitní.
- 9.25. Minimalizujte $\|\mathbf{C} \mathbf{x}\|_2$ za podmínky $\mathbf{A} \mathbf{x} = \mathbf{b}$, kde \mathbf{A} má lineárně nezávislé řádky a \mathbf{C} má lineárně nezávislé sloupce.
- 9.26. (★) Minimalizujte $\|\mathbf{A} \mathbf{x} - \mathbf{b}\|_2$ za podmínky $\mathbf{C} \mathbf{x} = \mathbf{0}$, kde \mathbf{A} má lineárně nezávislé sloupce a \mathbf{C} má lineárně nezávislé řádky.
- 9.27. (★) Minimalizujte $\|\mathbf{C} \mathbf{x}\|_2$ za podmínek $\mathbf{A} \mathbf{x} = \mathbf{0}$ a $\mathbf{x}^T \mathbf{x} = 1$.
- 9.28. (★) Minimalizujte $\|\mathbf{A} \mathbf{x}\|_2$ za podmínky $\mathbf{x}^T \mathbf{C} \mathbf{x} = 1$, kde \mathbf{C} je pozitivně definitní.
- 9.29. (★) Minimalizujte $\mathbf{a}^T \mathbf{x}$ za podmínky $\mathbf{x}^T \mathbf{C} \mathbf{x} = 1$, kde \mathbf{C} je pozitivně definitní.
- 9.30. (★) Jaké musí být vlastnosti matice \mathbf{A} a vektoru \mathbf{b} , aby $\max\{\|\mathbf{A} \mathbf{x}\|_2 \mid \mathbf{b}^T \mathbf{x} = 0\} = 0$?
- 9.31. Najděte (globální) minimum funkce $f(x, y) = x^2 y + y^2 + x$ za podmínek $-1 \leq x - y \leq 1$.

Nápověda ke cvičením

- 9.5.d) Stacionární body jsou 4.
- 9.5.e) Stacionární body jsou 3.
- 9.5.f) Stacionárních bodů je 5.
- 9.5.g) Stacionární body jsou 3, a to $(0, 0, 0)$, $(3/2, 3/2, -9/4)$, $(3/2, 3/2, -9/4)$.
- 9.25. $\mathbf{x} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T (\mathbf{A} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{A}^T)^{-1} \mathbf{b}$.
- 9.26. $\mathbf{x} = [\mathbf{I} - (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{C}^T (\mathbf{C} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{C}^T)^{-1} \mathbf{C}^T] (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$

Další příklady i s řešeními jsou v Ptákových skriptech Calculus.

Kapitola 10

Iterační algoritmy na volné lokální extrémů

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

Úvod: proč potřebujeme hledat lokální minima, najít lokální minimum je snadnější než najít řešení soustavy nelineárních rovnic. Tato kapitola jde říct za jednu přednášku, ale musí se hodně spěchat.

10.1 Rychlost konvergence iteračních algoritmů

Iterační algoritmus na řešení nějaké úlohy konstruuje posloupnost bodů \mathbf{x}_k , která konverguje k řešení úlohy \mathbf{x} . Posloupnost zbytků $r_k = \|\mathbf{x}_k - \mathbf{x}\|$ je nezáporná, $r_k \geq 0$, a konverguje k nule, $\lim_{k \rightarrow \infty} r_k = 0$. Zkoumejme rychlost konvergence této posloupnosti.

Pokud existuje limita

$$\lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k} = \rho, \quad (10.1)$$

řekneme, že posloupnost $\{r_k\}$ konverguje

- **sublineárně**, pokud $\rho = 1$,
- **lineárně**, pokud $0 < \rho < 1$,
- **superlineárně**, pokud $\rho = 0$.

Je jasné, že čím je ρ menší, tím posloupnost konverguje ‘rychleji’. Sublineární konvergence znamená velmi (často nepoužitelně) pomalý algoritmus. Lineární konvergence znamená přijatelnou rychlost, přibližně rovnou rychlosti konvergence geometrické řady. Mnoho široce používaných iteračních algoritmů konverguje lineárně. Superlineární konvergence znamená výtečný algoritmus.

Příklad 10.1.

1. Posloupnost $\{r_k\} = \{2^{-k}\} = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots\}$ konverguje lineárně, protože $r_{k+1}/r_k = 1/2$, což je nezávislé na k . Posloupnost je geometrická řada.
2. Posloupnost $\{r_k\} = \{1/k\} = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$ konverguje sublineárně, protože $r_{k+1}/r_k = k/(k+1)$, což pro $k \rightarrow \infty$ se blíží 1.

3. Posloupnost $\{r_k\} = \{2^{-2^k}\} = \{\frac{1}{4}, \frac{1}{16}, \frac{1}{256}, \dots\}$ konverguje superlineárně, protože

$$\frac{r_{k+1}}{r_k} = \frac{2^{-2^{k+1}}}{2^{-2^k}} = 2^{-2^{k+1}+2^k} = 2^{-2^k}$$

a tedy limita (10.1) je rovna 0.

Uvědomte si, jak fantasticky rychlá je to konvergence. Znamená to, že $r_{k+1} = r_k^2$, tj. s každou iterací se zhruba *zdvojnásobí počet platných cifer*. Strojové přesnosti dosáhneme za několik málo iterací.

4. Posloupnost $\{r_k\} = \{k^{-k}\}$ konverguje superlineárně (limitu (10.1) spočtete jako cvičení!).
5. Pro posloupnost $\{r_k\} = \{\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \dots\}$, tj. ‘koktavou’ verzi posloupnosti $\{2^{-k}\}$, limita (10.1) neexistuje, protože výraz r_{k+1}/r_k je jiný pro sudé a liché k .
6. Pro posloupnost $\{r_k\} = \{1, \frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{3}, \frac{1}{8}, \frac{1}{4}, \frac{1}{16}, \dots\}$, tj. proložené posloupnosti $\{1/k\}$ a $\{2^{-k}\}$, limita (10.1) neexistuje z podobného důvodu. \square

Poslední dva příklady ukazují nedostatečnost stávající definice: limita (10.1) neexistuje, přestože posloupnost je jinak ‘rozumná’. Proto se zavádí obecnější definice. Řekneme, že posloupnost $\{r_k\}$ konverguje **alespoň** sublineárně [lineárně, superlinárně], existuje-li posloupnost $\{a'_k\}$, která konverguje sublineárně [lineárně, superlinárně] a $a'_k \geq r_k$ pro každé k . Např. posloupnost z příkladu 5 výše konverguje alespoň lineárně, protože můžeme zvolit $a'_k = 2^{-k/2}$.

10.2 (★) Metoda zlatého řezu

Půlení intervalu je známá iterační metoda na hledání nulové hodnoty spojitě funkce $g: \mathbb{R} \rightarrow \mathbb{R}$ (tj. hledání kořene rovnice $g(x) = 0$). Metoda nevyžaduje počítání derivací funkce, které ani nemusí existovat. Na začátku k -té iterace máme dva body $x_1 < x_2$ takové, že

$$g(x_1)g(x_2) < 0. \tag{10.2}$$

To zaručuje, že v intervalu $[x_1, x_2]$ leží aspoň jeden kořen. V $(k+1)$ -ní iteraci přidáme bod $x_3 = (x_1 + x_2)/2$. Nezbytně bude buď $g(x_1)g(x_3) < 0$ nebo $g(x_3)g(x_2) < 0$ nebo $g(x_3) = 0$. V prvním případě interval $[x_1, x_2]$ nahradíme intervalem $[x_1, x_3]$, ve druhém případě intervalem $[x_3, x_2]$. Pokračujeme stejně. Protože v každé iteraci se interval neúčitosti zúží na polovinu, metoda konverguje lineárně ($r = 1$).

Hledejme nyní nikoliv nulovou hodnotu, ale minimum spojitě funkce $f: \mathbb{R} \rightarrow \mathbb{R}$. Funkci nazveme **unimodální** na intervalu $[x_1, x_2]$, pokud existuje bod x takový, že $x_1 < x < x_2$ a na intervalu $[x_1, x]$ funkce striktně klesá a na intervalu $[x, x_2]$ striktně roste. V tom případě má funkce na intervalu právě jedno minimum x , které se nabývá v jeho vnitřním bodě.

Zatímco pro zachycení kořene stačila dvojice bodů splňující (10.2), pro zachycení minima potřebujeme *trojici* bodů. Necht' v k -té iteraci máme tři body $x_1 < x_3 < x_2$ tak, že funkce je na intervalu $[x_1, x_2]$ unimodální a platí

$$(f(x_3) - f(x_1))(f(x_2) - f(x_3)) < 0. \tag{10.3}$$

Trojici (x_1, x_3, x_2) říkáme **závorka** (*bracket*). V $(k+1)$ -ní přidáme bod x_4 , dejme tomu mezi body x_3 a x_2 . Musí nastat jeden z těchto případů:

- $f(x_3) \leq f(x_4)$: funkce je unimodální na intervalu $[x_1, x_3]$ a minimum je zachyceno závorkou (x_1, x_3, x_4) .
- $f(x_3) > f(x_4)$: funkce je unimodální na intervalu $[x_3, x_2]$ a minimum je zachyceno závorkou (x_3, x_4, x_2) .

Zůstává otázka, jak volit pozici bodů, aby bylo zaručeno největší možné zmenšení intervalu neurčitosti, a to při obou možnostech 1 a 2. Chceme, aby závorky (x_1, x_3, x_2) , (x_1, x_3, x_4) a (x_3, x_4, x_2) byly rozděleny ve stejném poměru. Tedy

$$\frac{b}{a} = \frac{a}{c} = \frac{b-c}{c},$$

kde $a = x_3 - x_1$, $b = x_2 - x_3$, $c = x_4 - x_3$. Odtud dostaneme $\varphi - \varphi^{-1} = 1$, kde jsme označili $b/a = \varphi$. Kladný kořen této rovnice je číslo $\varphi = (1 + \sqrt{5})/2 \approx 1.618$, známé z antiky jako **zlatý řez**. Máme zaručeno, že v další iteraci bude interval neurčitosti φ -krát kratší.

Protože se v každé iteraci interval neurčitosti zmenší φ -krát, algoritmus konverguje lineárně ($r = \varphi^{-1}$).

Kdy algoritmus půlení intervalu a algoritmus zlatého řezu ukončit? Lze ukázat, že kvůli zaokrouhlovacím chybám nejde interval neurčitosti zmenšit na méně než asi $\sqrt{\varepsilon}$, kde ε je strojová přesnost.

10.3 Sestupné metody

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

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{v}_k, \quad (10.4)$$

kde vektor $\mathbf{v}_k \in \mathbb{R}^n$ je **směr hledání** a skalár $\alpha_k > 0$ je **délka kroku**. Ve třídě algoritmů zvaných **sestupné metody** (*descent methods*) hodnota účelové funkce monotonně klesá¹, $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$.

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

$$f'(\mathbf{x}_k) \mathbf{v}_k < 0, \quad (10.5)$$

tedy směrová derivace ve směru \mathbf{v}_k je záporná. Pokud v bodě \mathbf{x}_k existuje sestupný směr, existuje délka kroku $\alpha_k > 0$ tak, že $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$. Pokud v bodě \mathbf{x}_k sestupný směr neexistuje, vektor $f'(\mathbf{x}_k)$ je nutně nulový (proč?) a tedy \mathbf{x}_k je stacionární bod.

Máme-li sestupný směr, optimální délku kroku α_k najdeme minimalizací funkce f na polo-přímce z bodu \mathbf{x}_k ve směru \mathbf{v}_k . Tedy minimalizujeme funkci jedné proměnné

$$\varphi(\alpha_k) = f(\mathbf{x}_k + \alpha_k \mathbf{v}_k) \quad (10.6)$$

přes všechny $\alpha_k \geq 0$. Tato úloha je v kontextu vícerozměrné optimalizace nazývána *line search*. Úlohu stačí řešit přibližně. Takovou přibližnou metodu není obtížné vymyslet a proto se jí dále nebudeme zabývat. Poznamenejme ale, že metodu zlatého řezu nelze beze změn použít, protože funkce φ nemusí být unimodální.

Dále uvedeme nejznámější zástupce sestupných metod.

¹ Existují totiž i algoritmy, ve kterých hodnota $f(\mathbf{x}_k)$ neklesá monotonně (tj. někdy stoupne a někdy klesne) a přesto konvergují k optimu (např. *subgradientní metody*).

10.4 Gradientní metoda

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

$$\mathbf{v}_k = -f'(\mathbf{x}_k)^T = -\nabla f(\mathbf{x}_k). \quad (10.7)$$

Tento směr je sestupný, což je okamžitě vidět dosazením do (10.5).

Rychlost konvergence gradientní metody je lineární. Konvergence je často pomalá kvůli ‘cik-cak’ chování. Výhodou metody je její spolehlivost, protože směr je vždy sestupný.

10.4.1 (★) Závislost na lineární transformaci souřadnic

Transformujme vektor proměnných \mathbf{x} lineární transformací $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x}$, kde \mathbf{A} je čtvercová regulární matice. Je jasné, že úloha v nových proměnných bude mít stejné optimum jako v původních proměnných. Tedy

$$\min_{\mathbf{x}} f(\mathbf{x}) = \min_{\tilde{\mathbf{x}}} \tilde{f}(\tilde{\mathbf{x}}), \quad \text{kde} \quad \tilde{f}(\tilde{\mathbf{x}}) = \tilde{f}(\mathbf{A}\mathbf{x}) = f(\mathbf{x}) = f(\mathbf{A}^{-1}\tilde{\mathbf{x}}).$$

Iterace gradientní metody v nových proměnných je

$$\tilde{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_k - \alpha_k \tilde{f}'(\tilde{\mathbf{x}}_k)^T. \quad (10.8)$$

Zkoumejme, jaké iteraci to odpovídá v původních proměnných. K tomu potřebujeme vyjádřit (10.8) v proměnných \mathbf{x} . Použitím řetězového pravidla odvodíme

$$\tilde{f}'(\tilde{\mathbf{x}}) = \frac{d\tilde{f}(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}} = \frac{d\tilde{f}(\tilde{\mathbf{x}})}{d\mathbf{x}} \frac{d\mathbf{x}}{d\tilde{\mathbf{x}}} = \frac{df(\mathbf{x})}{d\mathbf{x}} \frac{d\mathbf{x}}{d\tilde{\mathbf{x}}} = f'(\mathbf{x})\mathbf{A}^{-1}.$$

Dosazením za $\tilde{\mathbf{x}}$ a $\tilde{f}'(\tilde{\mathbf{x}})$ do (10.8) a úpravou dostaneme

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k (\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (10.9)$$

To lze napsat ve tvaru (10.4) se směrem hledání

$$\mathbf{v}_k = -(\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T. \quad (10.10)$$

Tento směr se liší od původního směru (10.7) vynásobením maticí $(\mathbf{A}^T \mathbf{A})^{-1}$. Vidíme tedy, že gradientní metoda *není invariantní* vůči lineární transformaci souřadnic.

Ovšem lze ukázat, že nový směr (10.10) je také sestupný. Dosazením (10.7) do (10.5) to znamená, že $-f'(\mathbf{x}_k)(\mathbf{A}^T \mathbf{A})^{-1} f'(\mathbf{x}_k)^T < 0$. To je ale pravda, neboť matice $\mathbf{A}^T \mathbf{A}$ a tedy i její inverze je pozitivně definitní, viz Cvičení 6.16.

Na vzorec (10.10) 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 pozitivně definitní. Dá se ukázat i opak, totiž že každý sestupný směr lze napsat takto. Matice \mathbf{C}_k může být jiná v každém kroku. Uvidíme, že algoritmy uvedené dále budou mít vždy tento tvar.

10.5 Newtonova metoda

Newtonova metoda (přesněji Newton-Raphsonova) je slavný iterační algoritmus na řešení soustav nelineárních rovnic. Lze ho použít i na minimalizaci funkce tak, že hledáme nulový gradient. Oba způsoby použití popíšeme.

10.5.1 Použití na soustavy nelineárních rovnic

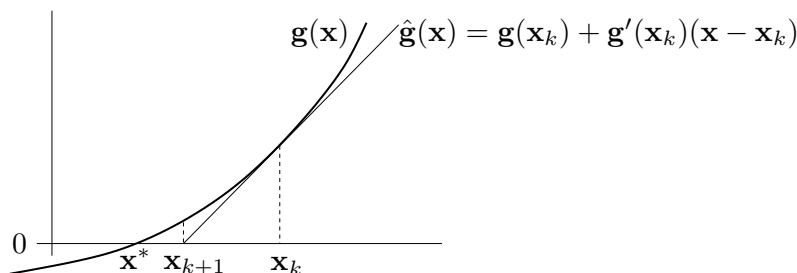
Řešme rovnici $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ je diferencovatelné zobrazení. Jedná se tedy o soustavu n rovnic s n neznámými. Tyto rovnice obecně mohou být nelineární. Zobrazení \mathbf{g} aproximujeme v okolí bodu \mathbf{x}_k Taylorovým polynomem prvního stupně

$$\mathbf{g}(\mathbf{x}) \approx \hat{\mathbf{g}}(\mathbf{x}) = \mathbf{g}(\mathbf{x}_k) + \mathbf{g}'(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k), \quad (10.11)$$

kde Jacobiho matice $\mathbf{g}'(\mathbf{x}_k) \in \mathbb{R}^{n \times n}$ je derivace zobrazení v bodě \mathbf{x}_k . Další iteraci \mathbf{x}_{k+1} najdeme řešením nehomogenní lineární soustavy $\hat{\mathbf{g}}(\mathbf{x}_{k+1}) = \mathbf{0}$. Pokud je Jacobiho matice regulární, řešením je

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^{-1} \mathbf{g}(\mathbf{x}_k). \quad (10.12)$$

Viz obrázek:



Lze dokázat, Newtonova metoda konverguje obvykle (i když ne vždy) superlineárně, tedy velmi rychle. Její nevýhodou je, že je nutno začít poměrně přesnou aproximací \mathbf{x}_0 skutečného řešení, jinak algoritmus snadno diverguje.

Příklad 10.2. *Babylónská metoda* na výpočet druhé odmocniny čísla $a \geq 0$ je dána iterací

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

To není nic jiného než Newtonova metoda pro řešení rovnice $0 = g(x) = x^2 - a$. Opravdu,

$$x_{k+1} = x_k - \frac{g(x)}{g'(x)} = x_k - \frac{x^2 - a}{2x} = x_k - \frac{1}{2} \left(x_k - \frac{a}{x_k} \right) = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right). \quad \square$$

Příklad 10.3. Hledejme průsečík křivek $(x-1)^2 + y^2 = 1$ a $x^4 + y^4 = 1$. Máme $n = 2$ a

$$\mathbf{x} = (x, y) = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{g}(\mathbf{x}) = \mathbf{g}(x, y) = \begin{bmatrix} (x-1)^2 + y^2 - 1 \\ x^4 + y^4 - 1 \end{bmatrix}, \quad \mathbf{g}'(\mathbf{x}) = \mathbf{g}'(x, y) = \begin{bmatrix} 2(x-1) & 2y \\ 4x^3 & 4y^3 \end{bmatrix}.$$

Iterace (10.12) je

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

Načrtneme-li si obě křivky, vidíme, že mají dva průsečíky. Zvolme počáteční odhad pro horní průsečík $(x_0, y_0) = (1, 1)$. První iterace bude

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

Šestá iterace $(x_6, y_6) = (0.671859751039018, 0.944629015546222)$ je taková, že rovnice jsou splněny se strojovou přesností. \square

Příklad 10.4. Funkce $g(x) = x^2 - 1$ má dva nulové body $x = \pm 1$. Pokud v nějaké iteraci bude $x_k = 0$, nastane dělení nulou. Pokud bude x_k velmi malé, dělení nulou nenastane, ale iterace x_{k+1} se ocitne velmi daleko od kořene. \square

Příklad 10.5. Pro funkci $g(x) = x^3 - 2x + 2$ zvolme $x_0 = 0$. Další iterace bude $x_1 = 1$ a další $x_2 = 0$. Algoritmus bude oscilovat mezi hodnotami 0 a 1, tedy bude divergovat. \square

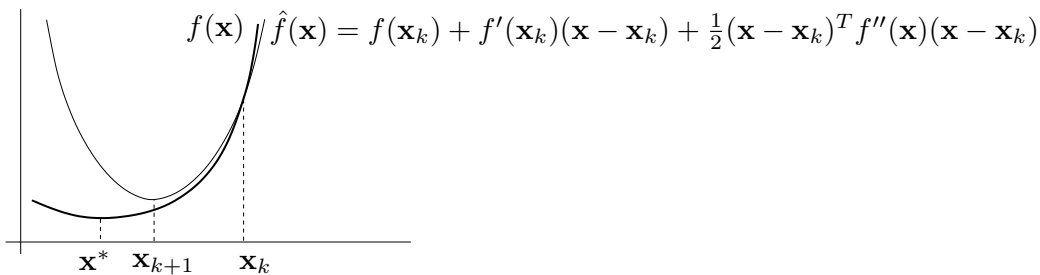
10.5.2 Použití na minimalizaci funkce

Newtonovu metodu lze použít pro hledání lokálního extrému dvakrát diferencovatelné funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tak, že v algoritmu (10.12) položíme $\mathbf{g}(\mathbf{x}) = f'(\mathbf{x})^T$. Tím dostaneme iteraci

$$\mathbf{x}_{k+1} = \mathbf{x}_k - f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T, \quad (10.13)$$

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

Význam iterace (10.12) byl takový, že se zobrazení \mathbf{g} aproximovalo Taylorovým polynomem prvního stupně (tedy afinním zobrazením) a pak se našel kořen tohoto polynomu. Význam iterace (10.13) je takový, že se funkce f aproximuje Taylorovým polynomem druhého stupně (tedy kvadratickou funkcí) a pak se najde minimum této kvadratické funkce. Odvoďte podrobně, že tomu tak je!



Iteraci (10.13) lze napsat v obecnějším tvaru (10.4), kde

$$\mathbf{v}_k = -f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T. \quad (10.14)$$

Výhodou tohoto zobecnění je možnost zvolit optimální (ne nutně jednotkovou) délku kroku pomocí jednorozměrné minimalizace (10.6). Algoritmu (10.13) s jednotkovou délkou kroku se pak říká **čistá** Newtonova metoda.

Vektoru (10.14) říkáme **Newtonův směr**. Aby to byl sestupný směr, musí být

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

Postačující podmínkou pro to je, aby matice $f''(\mathbf{x}_k)$ byla pozitivně definitní (neboť pak bude pozitivně definitní i její inverze, viz Cvičení 6.18).

10.6 Gauss-Newtonova metoda

Řešme přeурčenou soustavu rovnic $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ pro $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (tedy soustavu m rovnic s n neznámými) ve smyslu nejmenších čtverců. To vede na minimalizaci funkce

$$f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|_2^2 = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x}) = \sum_{i=1}^m g_i(\mathbf{x})^2, \quad (10.15)$$

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

Všimněte si, že zatímco v §10.4 a §10.5.2 bylo cílem minimalizovat *obecnou* funkci, zde chceme minimalizovat funkci ve speciálním tvaru (10.15). Nyní máme dvě možnosti. Buď můžeme nasadit na funkci (10.15) jednu z metod pro minimalizaci obecné funkce, k čemuž se vrátíme v §10.6.1. Nebo můžeme být chytřejší a využít speciálního tvaru funkce (10.15), což uděláme teď.

Aproximujme opět zobrazení \mathbf{g} Taylorovým polynomem prvního stupně (10.11). Úloha (10.15) pak vyžaduje minimalizovat $\|\hat{\mathbf{g}}(\mathbf{x})\|_2^2$. To je úloha lineárních nejmenších čtverců, kterou již známe z §5.1. Vede na normální rovnici

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

Pokud má Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ plnou hodnotu, tuto rovnici můžeme vyřešit pomocí pseudo-inverze:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \underbrace{(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T}_{\mathbf{g}'(\mathbf{x}_k)^+} \mathbf{g}(\mathbf{x}_k) \quad (10.16)$$

Algoritmus (10.16) je znám jako **Gauss-Newtonova metoda**. Můžeme jej opět napsat obecněji ve tvaru (10.4) se směrem hledání

$$\mathbf{v}_k = -(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}(\mathbf{x}_k). \quad (10.17)$$

Pro $m = n$ máme $\mathbf{g}'(\mathbf{x}_k)^+ = \mathbf{g}'(\mathbf{x}_k)^{-1}$, tedy Gauss-Newtonova metoda se redukuje na Newtonovu metodu (10.12) na řešení soustavy n rovnic s n neznámými.

Snadno spočítáme (viz §8.3.2) derivaci účelové funkce (10.15), je rovna $f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$. Z toho vidíme, že Gauss-Newtonův směr (10.17) lze psát ekvivalentně jako

$$\mathbf{v}_k = -\frac{1}{2}(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1} f'(\mathbf{x}_k)^T. \quad (10.18)$$

Tento směr se liší od gradientního směru (10.7) pouze násobením maticí $\frac{1}{2}(\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k))^{-1}$. Pokud Jacobián $\mathbf{g}'(\mathbf{x}_k)$ má plnou hodnotu (tedy n), tato matice je pozitivně definitní. Podobnou úvahou jako v §10.5.2 dostaneme, že směr (10.17) je *vždy* sestupný.

Lze dokázat, že 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 (10.15). Protože ale Gauss-Newtonův směr je vždy sestupný, vhodnou volbou délky kroku α_k lze vždy zajistit konvergenci.

Příklad 10.6. V systému GPS máme m satelitů se známými souřadnicemi $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a chceme spočítat souřadnice pozorovatele $\mathbf{x} \in \mathbb{R}^n$ z naměřených vzdáleností $y_i = \|\mathbf{a}_i - \mathbf{x}\|_2$ pozorovatele od satelitů. Měření jsou zatížena chybou, proto obecně tato soustava rovnic nebude mít žádné řešení. Řešme tuto přeurčenou nelineární soustavu ve smyslu nejmenších čtverců, tedy minimalizujme funkci

$$f(\mathbf{x}) = \sum_{i=1}^m (\|\mathbf{x} - \mathbf{a}_i\|_2 - y_i)^2.$$

Máme tedy $\mathbf{g} = (g_1, \dots, g_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$, kde $g_i(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}_i\|_2 - y_i$. Derivace složek \mathbf{g} je (pomůžte nám §8.3.2, ale udělejte sami!) $g'_i(\mathbf{x}) = (\mathbf{x} - \mathbf{a}_i)^T / \|\mathbf{x} - \mathbf{a}_i\|_2$. Tedy

$$\mathbf{g}'(\mathbf{x}) = \begin{bmatrix} (\mathbf{x} - \mathbf{a}_1)^T / \|\mathbf{x} - \mathbf{a}_1\|_2 \\ \vdots \\ (\mathbf{x} - \mathbf{a}_m)^T / \|\mathbf{x} - \mathbf{a}_m\|_2 \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Pak dosadíme do vzorečku (10.16). □

10.6.1 Rozdíl proti Newtonově metodě

Předpokládejme, že bychom optimalizovali naši účelovou funkci (10.15) přímo Newtonovou metodou z §10.5.2. Spočítejme (proved'te sami!) Hessián funkce (10.15):

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

Hessián je součtem členu obsahujícího derivace prvního řádu a členu obsahujícího derivace druhého řádu. Vidíme, že směr (10.18) se liší od Newtonova směru (10.14) zanedbáním členu druhého řádu v Hessiánu $f''(\mathbf{x}_k)$. To se projevuje tím, že Gauss-Newtonova metoda má horší lokální konvergenční chování než plná Newtonova metoda – ani v blízkém okolí řešení nemusí konvergovat superlineárně. Na druhou stranu, vyhnuli jsme se počítání druhých derivací funkce \mathbf{g} , což je velké zjednodušení.

10.6.2 Levenberg-Marquardtova metoda

Levenberg-Marquardtova metoda je široce používané vylepšení Gauss-Newtonovy metody, které matici $\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x})$ v iteraci (10.16) nahrazuje maticí

$$\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k) + \mu_k \mathbf{I} \tag{10.19}$$

pro nějaké zvolené $\mu_k > 0$. Vidíme, že:

- Pro malé μ_k se Levenberg-Marquardtova iterace blíží Gauss-Newtonově iteraci.
- Pro velké μ_k je inverze matice (10.19) blízká $\mu_k^{-1} \mathbf{I}$, tedy Levenberg-Marquardtova iterace je blízká $\mathbf{x}_{k+1} = \mathbf{x}_k - \mu_k^{-1} f'(\mathbf{x}_k)^T$. Ale to je iterace gradientní metody s délkou kroku μ_k^{-1} .

Tím jsou spojeny výhody Gauss-Newtonovy metody (typicky rychlá konvergence v okolí optima) a gradientní metody (spolehlivost i daleko od optima). Volbou parametru μ_k spojitě přecházíme mezi oběma metodami.

Parametr μ_k měníme během algoritmu. Začneme např. s $\mu_0 = 10^3$ a pak v každé iteraci:

- Pokud iterace snížila účelovou funkci, iteraci přijmeme a μ_k zmenšíme.
- Pokud iterace nesnížila účelovou funkci, iteraci odmítneme a μ_k zvětšíme.

Zvětšování a zmenšování μ_k děláme násobením a dělením konstantou, např. 10. Všimněte si, toto nahrazuje optimalizaci délky kroku α_k (*line search*).

Na algoritmus lze pohlížet i jinak. V iteraci (10.16) se počítá inverze matice $\mathbf{g}'(\mathbf{x}_k)^T \mathbf{g}'(\mathbf{x}_k)$. Tato matice je sice vždy pozitivně semidefinitní, ale může být blízká singularní (kdy se to stane?). To neblaze ovlivní stabilitu algoritmu. Matice (10.19) je ale vždy pozitivně definitní (viz Cvičení 6.17), a tedy regulární.

10.7 Statistické odůvodnění kritéria nejmenších čtverců

Odhadujeme skryté parametry \mathbf{x} nějakého systému z měření \mathbf{y} na systému. Budiž vázány známou závislostí $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Měření jsou zatížena chybami, které jsou způsobeny šumem senzorů, nepřesnostmi měření, nedokonalou znalostí modelu, apod. Tedy

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}, \quad (10.20)$$

kde $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m)$ jsou náhodné proměnné modelující chyby měření $\mathbf{y} = (y_1, \dots, y_m)$. Metoda nejmenších čtverců říká, že máme minimalizovat $\|\boldsymbol{\varepsilon}\|_2^2 = \sum_{i=1}^m \varepsilon_i^2$, ale neříká proč.

Důvod odvodíme statistickou úvahou. Metoda činí dva předpoklady:

- Náhodné proměnné ε_i mají normální (neboli Gaussovo) rozdělení s nulovou střední hodnotou a směrodatnou odchylkou σ ,

$$p(\varepsilon_i) = c e^{-\varepsilon_i^2/(2\sigma^2)},$$

kde $c = (\sigma\sqrt{2\pi})^{-1}$ je normalizační konstanta.

- Náhodné proměnné $\varepsilon_1, \dots, \varepsilon_m$ jsou na sobě nezávislé. Tedy sdružená pravděpodobnost je rovna součinu

$$p(\boldsymbol{\varepsilon}) = p(\varepsilon_1, \dots, \varepsilon_m) = \prod_{i=1}^m p(\varepsilon_i) = \prod_{i=1}^m c e^{-\varepsilon_i^2/(2\sigma^2)}. \quad (10.21)$$

Dále použijeme *princip maxima věrohodnosti*. Ten říká, že parametry \mathbf{x} se mají najít tak, aby $p(\boldsymbol{\varepsilon}) = p(\mathbf{y} - \mathbf{f}(\mathbf{x}))$ bylo maximální. Je pohodlnější minimalizovat záporný logaritmus

$$-\log p(\varepsilon_1, \dots, \varepsilon_m) = -\sum_{i=1}^m \log p(\varepsilon_i) = \sum_{i=1}^m \left(\frac{\varepsilon_i^2}{2\sigma^2} - \log c \right).$$

Jelikož σ je konstanta, je to totéž jako minimalizovat $\sum_i \varepsilon_i^2$.

10.8 Cvičení

- 10.1. Najděte lokální extrém funkce $f(x, y) = x^2 - y + \sin(y^2 - 2x)$ čistou Newtonovou metodou. Počáteční odhad zvolte $(x_0, y_0) = (1, 1)$.
- 10.2. Máme m bodů v rovině o souřadnicích (x_i, y_i) , $i = 1, \dots, m$. Tyto body chceme proložit kružnicí ve smyslu nejmenších čtverců – tj. hledáme kružnici se středem (u, v) a poloměrem r takovou, aby součet čtverců kolmých vzdáleností bodů ke kružnici byl minimální. Zformulujte příslušnou optimalizační úlohu. Odvoďte iteraci Gauss-Newtonovy a Levenberg-Marquardtovy metody.

Kapitola 11

Konvexita

11.1 Konvexní množiny

Definice 11.1. Množina $X \subseteq \mathbb{R}^n$ se nazývá **konvexní**, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies \alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \in X. \quad (11.1)$$

Množina $\{\alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \mid 0 \leq \alpha \leq 1\}$ je úsečka spojující body \mathbf{x} a \mathbf{y} (viz Příklad 3.6). Definice tedy říká, že množina je konvexní, jestliže s každými dvěma body obsahuje i úsečku, která je spojuje. Obrázek ukazuje příklad konvexní a nekonvexní množiny v \mathbb{R}^2 :



Konvexní množinu lze definovat i abstraktněji. **Konvexní kombinace** vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je jejich lineární kombinace $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ taková, že $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Množina je konvexní právě tehdy, když je uzavřená vůči konvexním kombinacím (neboli každá konvexní kombinace vektorů z množiny leží v množině). Lze dokázat indukcí, že tato definice je ekvivalentní Definici 11.1. Všimněte si, že $\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}$ pro $0 \leq \alpha \leq 1$ je konvexní kombinací dvou vektorů \mathbf{x}, \mathbf{y} .

Konvexní obal vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich konvexních kombinací. Tuto k -tici vektorů můžeme vnímat jako množinu $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$, konvexní obal pak značíme

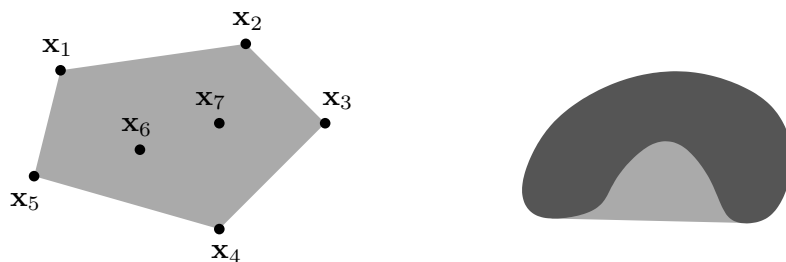
$$\text{conv } X = \text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} = \{\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k \mid \alpha_1 + \dots + \alpha_k = 1, \alpha_1, \dots, \alpha_k \geq 0\}. \quad (11.2)$$

Jak se definuje konvexní obal množiny s *nekonečným* počtem prvků, např. pravém obrázku výše? Nelze použít definice (11.2), neboť není jasné, co znamená součet $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ pro nekonečný počet vektorů (množina X může být i nespočetná). Konvexní obal libovolné (konečné či nekonečné) množiny $X \subseteq \mathbb{R}^n$ se definuje jako

$$\text{conv } X = \bigcap_{\substack{Y \supset X \\ Y \text{ konvexní}}} Y$$

tedy průnik všech konvexních množin, které množinu obsahují.

Obrázek ukazuje konvexní obal konečné (vlevo) a nekonečné (vpravo) množiny pro $n = 2$:



11.1.1 Čtyři kombinace a čtyři obaly

Konvexní kombinace je lineární kombinace, jejíž koeficienty splňují omezení $\alpha_1 + \dots + \alpha_k = 1$ a $\alpha_1, \dots, \alpha_k \geq 0$. Všimněte si, že když vynecháme druhé omezení, dostaneme afinní kombinaci (viz §3.3). Podle toho, které ze dvou omezení vyžadujeme, dostaneme čtyři druhy kombinací. Udělejme si v nich nyní pořádek.

Vážený součet $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ se nazývá jejich

- lineární kombinace**, jestliže $\alpha_1, \dots, \alpha_k \in \mathbb{R}$.
- afinní kombinace**, jestliže $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, $\alpha_1 + \dots + \alpha_k = 1$.
- nezáporná kombinace**, jestliže $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, $\alpha_1, \dots, \alpha_k \geq 0$.
- konvexní kombinace**, jestliže $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, $\alpha_1 + \dots + \alpha_k = 1$, $\alpha_1, \dots, \alpha_k \geq 0$.

Množina, která je uzavřená vůči

- lineárním kombinacím, se nazývá **lineární podprostor**.
- afinním kombinacím, se nazývá **afinní podprostor**.
- nezáporným kombinacím, se nazývá **konvexní kužel**.
- konvexním kombinacím, se nazývá **konvexní množina**.

K tomu, co již znáte, přibyl pojem nezáporné kombinace a konvexního kuželu.

Lineární [afinní, nezáporný, konvexní] **obal** vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich lineárních [afinních, nezáporných, konvexních] kombinací. Obecněji, obal (konečné či nekonečné) množiny $X \subseteq \mathbb{R}^n$ je množina kombinací všech konečných podmnožin X . Ekvivalentně, lineární [afinní, nezáporný, konvexní] obal množiny $X \subseteq \mathbb{R}^n$ je nejmenší lineární podprostor [afinní podprostor, konvexní kužel, konvexní množina] obsahující množinu X .

Jako cvičení si nakreslete lineární, afinní, nezáporný a konvexní obal náhodně zvolených k vektorů v \mathbb{R}^n pro devět případů $k, n \in \{1, 2, 3\}$.

11.1.2 Operace zachovávající konvexitu množin

Jaké operace s konvexními množinami mají za výsledek opět konvexní množinu? Zdaleka nej-důležitější taková operace je průnik. Následující větu je snadné dokázat.

Věta 11.1. *Průnik (konečně či nekonečně mnoha) konvexních množin je konvexní množina.*

Důkaz. Uděláme jen pro dvě množiny, zobecnění na libovolná počet množin je očividné. Necht' $X, Y \subseteq \mathbb{R}^n$ jsou konvexní. Necht' $\mathbf{x}, \mathbf{y} \in X \cap Y$, tedy každý z bodů \mathbf{x}, \mathbf{y} je současně v množině X i v množině Y . Proto pro $0 \leq \alpha \leq 1$ bude bod $\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}$ také v množině X i Y , tedy bude v množině $X \cap Y$. \square

Sjednocení konvexních množin ale *nemusí* být konvexní množina.

11.2 Konvexní funkce

Definice 11.2. Funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je **konvexní** na množině $X \subseteq \mathbb{R}^n$, jestliže množina X je konvexní a platí

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}). \quad (11.3)$$

Funkce f je **konkávni** na množině X , jestliže je funkce $-f$ konvexní na množině X .

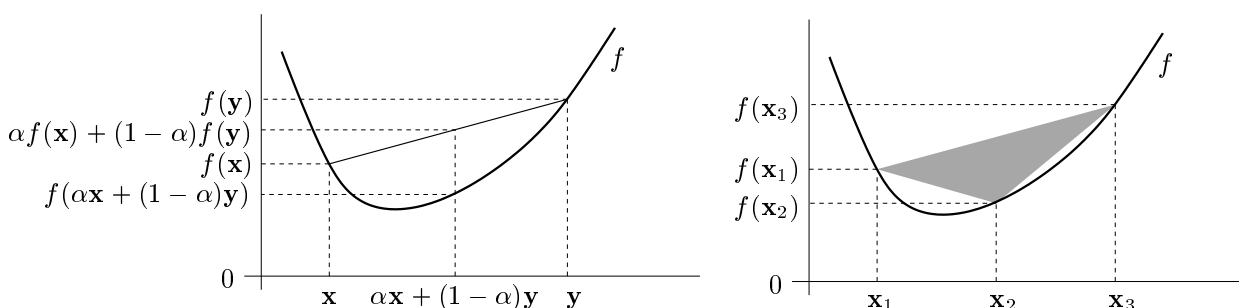
Rozlišujte pojem *konvexní množina* a *konvexní funkce*, jde o různé věci. Dále si všimněte, že X musí být konvexní množina – pojem konvexní funkce na nekonvexní množině nemá smysl. Pokud X je celý definiční obor funkce f , odkaz na X můžeme vynechat a říkáme pouze, že funkce f je konvexní.

Podmínku (11.3) lze zobecnit pro více než dva body: funkce f je konvexní právě tehdy, když

$$\mathbf{x}_1, \dots, \mathbf{x}_k \in X, \alpha_1, \dots, \alpha_k \geq 0, \alpha_1 + \dots + \alpha_k = 1 \implies f(\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k) \leq \alpha_1 f(\mathbf{x}_1) + \dots + \alpha_k f(\mathbf{x}_k), \quad (11.4)$$

neboli ‘funkční hodnota konvexní kombinace není větší než konvexní kombinace funkčních hodnot’. Podmínka (11.4) zjevně implikuje podmínku (11.3) a indukcí lze dokázat, že to platí i naopak. Podmínku (11.4) se někdy říká **Jensenova nerovnost**. Porovnejte ji s definicí lineárního zobrazení (3.3)!

Geometrický význam podmínky (11.3) 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 (11.4) je ten, že konvexní polyedr vybarvený šedě (viz pravý obrázek) leží nad grafem funkce. Podrobně rozmyslete, jak tyto geometrické interpretace odpovídají výrazům (11.3) a (11.4)!



Důkaz konvexity funkce z Definice 11.2 vyžaduje často intuici, neexistuje na to mechanický postup. Chceme-li dokázat, že funkce *není* konvexní, stačí nám nalézt jedinou trojici $(\mathbf{x}, \mathbf{y}, \alpha)$ porušující implikaci (11.3) – její ‘uhodnutí’ však vyžaduje, abychom měli o funkci představu.

Příklad 11.1. Dokažme z Definice 11.2, že funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ definovaná jako $f(\mathbf{x}) = \min_{i=1}^n x_i$ není konvexní. Např. volba $n = 2$, $\mathbf{x} = (0, 2)$, $\mathbf{y} = (2, 0)$, $\alpha = \frac{1}{2}$ nesplňuje (11.3), neboť

$$f((\mathbf{x} + \mathbf{y})/2) = f(1, 1) = 1 > (f(\mathbf{x}) + f(\mathbf{y}))/2 = (0 + 0)/2 = 0. \quad \square$$

Poznamenejme, že použitím Jensenovy nerovnosti na vhodnou konvexní funkci lze dokázat mnoho známých nerovností.

Příklad 11.2. Funkce \log je konkávní na \mathbb{R}_{++} . Napišme pro tuto funkci Jensenovu nerovnost (11.4) (jelikož funkce je konkávní a ne konvexní, musíme v Jensenově nerovnosti obrátit znaménko nerovnosti), ve které položíme $\alpha_1 = \dots = \alpha_n = \frac{1}{n}$:

$$\log \frac{x_1 + \dots + x_n}{n} \geq \frac{\log x_1 + \dots + \log x_n}{n}$$

kde x_1, \dots, x_n jsou kladné. Vezmeme-li exponenciálu každé strany, dostaneme

$$\frac{x_1 + \dots + x_n}{n} \geq (x_1 \dots x_n)^{1/n}$$

Tato známá nerovnost říká, že aritmetický průměr není nikdy menší než geometrický. □

Příklad 11.3. 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. Záporná entropie $f(x) = x \log x$ je konvexní na \mathbb{R}_{++} (nebo i na \mathbb{R}_+ , pokud dodefinujeme $0 \log 0 = 0$, což se často dělá, protože $\lim_{x \rightarrow 0^+} x \log x = 0$).
6. Afinní funkce $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$ je zároveň konvexní i konkávní.
7. Kvadratická forma $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ je konvexní pro \mathbf{A} pozitivně semidefinitní, konkávní pro \mathbf{A} negativně semidefinitní, a nekonvexní a nekonkávní pro \mathbf{A} indefinitní (viz Příklad 11.4).
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

$$f_a(\mathbf{x}) = f(a\mathbf{x})/a = \log(e^{ax_1} + \dots + e^{ax_n})/a$$

se pro $a \rightarrow +\infty$ blíží funkci $\max_{i=1}^n x_i$ (dokažte výpočtem limity!).

10. Geometrický průměr $f(\mathbf{x}) = (x_1 \dots x_n)^{1/n}$ je konkávní na \mathbb{R}_+^n .

Nakreslete či představte si vrstevnice a grafy těchto funkcí! □

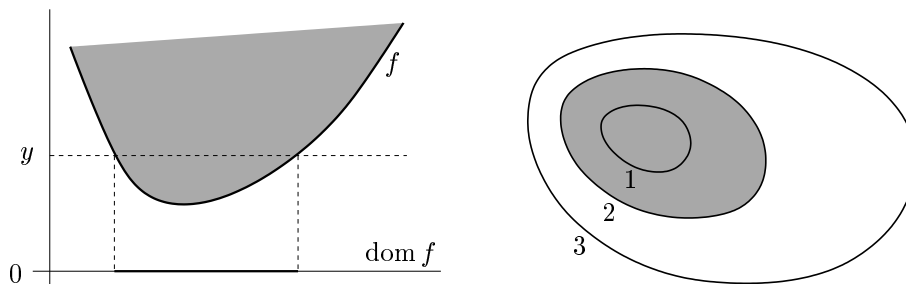
11.2.1 Epigraf a subkontura

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:

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

Levý obrázek znázorňuje subkonturu výšky y a epigraf funkce $\mathbb{R} \rightarrow \mathbb{R}$, pravý obrázek subkonturu výšky 2 funkce $\mathbb{R}^2 \rightarrow \mathbb{R}$:

¹ Slovo 'subkontura' je pokus o český překlad anglického 'sublevel set'.



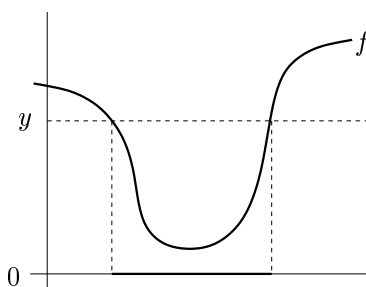
Věta 11.2. Je-li f konvexní funkce, pak je každá subkontura této funkce konvexní množina.

Důkaz. Předpokládejme, že body \mathbf{x}_1 a \mathbf{x}_2 patří do subkontury, tedy $f(\mathbf{x}_1) \leq y$ a $f(\mathbf{x}_2) \leq y$. Pro každé $0 \leq \alpha \leq 1$ platí

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \leq \alpha y + (1 - \alpha) y = y,$$

kde první nerovnost plyne z konvexity funkce a druhá z nerovností $f(\mathbf{x}_1) \leq y$, $f(\mathbf{x}_2) \leq y$. Tedy bod $\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$ patří do subkontury, která je proto konvexní množina. \square

Obrácená implikace ve Větě 11.2 neplatí: snadno najdeme funkci, která není konvexní a jejíž každá subkontura je konvexní množina². Příklad je na obrázku:



Věta 11.3. Funkce f je konvexní právě tehdy, když její epigraf je konvexní množina.

Důkaz. Předpokládejme, že funkce f je konvexní. Vezměme dva body (\mathbf{x}_1, y_1) a (\mathbf{x}_2, y_2) z epigrafu, tedy $f(\mathbf{x}_1) \leq y_1$ a $f(\mathbf{x}_2) \leq y_2$. Pro každé $0 \leq \alpha \leq 1$ platí

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \leq \alpha y_1 + (1 - \alpha) y_2,$$

kde první nerovnost plyne z konvexity funkce a druhá nerovnost z $f(\mathbf{x}_1) \leq y_1$ a $f(\mathbf{x}_2) \leq y_2$. Tedy bod $\alpha(\mathbf{x}_1, y_1) + (1 - \alpha)(\mathbf{x}_2, y_2)$ patří do epigrafu, který je proto konvexní množina.

Předpokládejme, že epigraf je konvexní množina. Tedy pokud body (\mathbf{x}_1, y_1) a (\mathbf{x}_2, y_2) patří do epigrafu, pak také bod $\alpha(\mathbf{x}_1, y_1) + (1 - \alpha)(\mathbf{x}_2, y_2)$ patří do epigrafu pro každé $0 \leq \alpha \leq 1$. Volbou $y_1 = f(\mathbf{x}_1)$ a $y_2 = f(\mathbf{x}_2)$ máme

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \alpha y_1 + (1 - \alpha) y_2 = \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2),$$

proto je funkce f konvexní. \square

² Funkce, jejíž každá subkontura je konvexní množina, se nazývá *kvazikonvexní*. Kvazikonvexní funkce nejsou zdaleka tak hezké jako konvexní funkce.

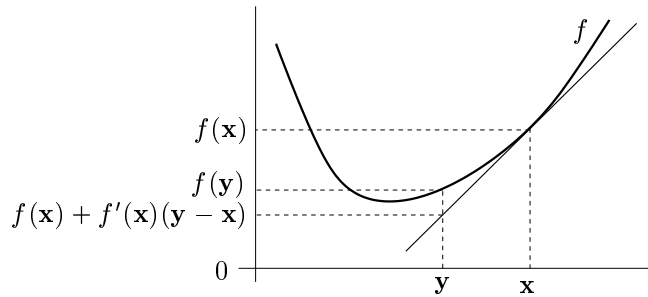
11.2.2 Konvexita diferencovatelných funkcí

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 Definice 11.2 charakterizovat pomocí derivací. Následující dvě věty uvedeme bez důkazů.

Věta 11.4. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je diferencovatelná na konvexní množině $X \subseteq \mathbb{R}^n$. Funkce f je konvexní na množině X právě tehdy, když

$$\mathbf{x} \in X, \mathbf{y} \in X \implies f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}).$$

To znamená, že tečný prostor ke grafu funkce v každém bodě $\mathbf{x} \in X$ leží celá (tj. pro každé \mathbf{y}) pod grafem (promyslete a porovnejte s definicí derivace v §8.3!):



Věta 11.5. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je dvakrát diferencovatelná na konvexní množině $X \subseteq \mathbb{R}^n$. Funkce f je konvexní na množině X právě tehdy, když v každém bodě $\mathbf{x} \in X$ je Hessova matice $f''(\mathbf{x})$ pozitivně semidefinitní.

Příklad 11.4. 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 11.5. To je triviální, protože Hessián je $f''(\mathbf{x}) = 2\mathbf{A}$ a tedy je pozitivně semidefinitní.
- Dokažme konvexitu z Věty 11.4. 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

$$\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y} = (\mathbf{x} - \mathbf{y})^T \mathbf{A}(\mathbf{x} - \mathbf{y}), \quad (11.5)$$

což je nezáporné pro každé \mathbf{x}, \mathbf{y} , protože \mathbf{A} je pozitivně semidefinitní.

- Dokažme konvexitu z Definice 11.2. Musíme dokázat, že pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $0 \leq \alpha \leq 1$ platí (11.3), tedy

$$[\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}]^T \mathbf{A}[\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}] \leq \alpha \mathbf{x}^T \mathbf{A} \mathbf{x} + (1 - \alpha)\mathbf{y}^T \mathbf{A} \mathbf{y}$$

Po roznásobení a převedení všech členů na jednu stranu upravujeme:

$$\begin{aligned} (\alpha - \alpha^2)\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\alpha(1 - \alpha)\mathbf{x}^T \mathbf{A} \mathbf{y} + ((1 - \alpha) - (1 - \alpha)^2)\mathbf{y}^T \mathbf{A} \mathbf{y} &\geq 0 \\ \alpha(1 - \alpha)(\mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{y}^T \mathbf{A} \mathbf{y}) &\geq 0. \end{aligned}$$

Výraz $\alpha(1 - \alpha)$ je pro každé $0 \leq \alpha \leq 1$ nezáporný. Nezápornost výrazu (11.5) jsme již ukázali. \square

³ Všimněte si, že pro $n = 1$ a $\mathbf{A} = 1$ se rovnost (11.5) zjednoduší na známé $x^2 - 2xy + y^2 = (x - y)^2$.

11.2.3 Operace zachovávající konvexitu funkcí

Operace zachovávající konvexitu funkcí umožňují z jednoduchých konvexních funkcí získat složitější. Konvexitu složitější funkce je často snadnější dokázat pohodlněji pomocí těchto operací než z Definice 11.2.

Jsou-li $g_1, \dots, g_k: \mathbb{R}^n \rightarrow \mathbb{R}$ konvexní funkce a $\alpha_1, \dots, \alpha_k \geq 0$, je snadné dokázat z Definice 11.2 (proved'te!), že také funkce

$$f = \alpha_1 g_1 + \dots + \alpha_k g_k$$

je konvexní. Speciálně, jsou-li f a g konvexní funkce, pak $f + g$ je konvexní.

Zkoumejme nyní složenou funkci $f(\mathbf{x}) = (h \circ \mathbf{g})(\mathbf{x}) = h(\mathbf{g}(\mathbf{x}))$, kde $\mathbb{R}^n \xrightarrow{\mathbf{g}} \mathbb{R}^m \xrightarrow{h} \mathbb{R}$. Obecně *neplatí* ani v případě $m = n = 1$, že konvexitu funkcí g a h zaručuje konvexitu funkce f . Nutné a postačující podmínky pro konvexitu složené funkce jsou obecně dosti komplikované a nebudeme je uvádět. Uvedeme jen nejdůležitější případ, kdy \mathbf{g} je afinní zobrazení.

Věta 11.6. *Necht' funkce $h: \mathbb{R}^m \rightarrow \mathbb{R}$ je konvexní. Necht' $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{b} \in \mathbb{R}^m$. Pak funkce $f(\mathbf{x}) = h(\mathbf{A}\mathbf{x} + \mathbf{b})$ je konvexní.*

Důkaz. Pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ a $0 \leq \alpha \leq 1$ platí

$$\begin{aligned} f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) &= h(\mathbf{A}[\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}] + \mathbf{b}) \\ &= h(\alpha(\mathbf{A}\mathbf{x} + \mathbf{b}) + (1 - \alpha)(\mathbf{A}\mathbf{y} + \mathbf{b})) \\ &\leq \alpha h(\mathbf{A}\mathbf{x} + \mathbf{b}) + (1 - \alpha)h(\mathbf{A}\mathbf{y} + \mathbf{b}) \\ &= \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}). \end{aligned} \quad \square$$

Nejzajímavější operace zachovávající konvexitu funkcí je maximum.

Věta 11.7. *Necht' I je libovolná množina a $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I$, jsou konvexní funkce. Pak funkce*

$$f(\mathbf{x}) = \max_{i \in I} g_i(\mathbf{x}) \tag{11.6}$$

je konvexní, kde předpokládáme, že pro každé \mathbf{x} maximum existuje⁴.

Důkaz. Protože funkce g_i jsou konvexní, dle Věty 11.3 jsou jejich epigrafy konvexní množiny. Snadno ověříme, že epigraf funkce (11.6) je průnik epigrafů funkcí g_i . Dle Věty 11.1 je průnik konvexních množin konvexní množina. Tedy epigraf funkce (11.6) je konvexní množina. Dle Věty 11.3 je tedy funkce f konvexní. \square

Velká obecnost věty plyne z toho, že indexová množina I může být konečná i nekonečná (a to spočetná i nespočetná). Uvedme nejprve příklady pro konečnou množinu I .

Příklad 11.5. Funkce

$$f(\mathbf{x}) = \max_{i=1}^k (\mathbf{a}_i^T \mathbf{x} + b_i)$$

je maximumm afinních funkcí. Protože afinní funkce jsou konvexní, je i jejich maximum konvexní. Tuto funkci uvádíme také v §12.3.1. \square

⁴ Pokud pro nějaké \mathbf{x} množina $\{g_i(\mathbf{x}) \mid i \in I\}$ nemá největší prvek (což se může stát jen tehdy, je-li množina I nekonečná), můžeme maximum v (11.6) nahradit supremem a věta stále platí.

Příklad 11.6. Necht' $f(\mathbf{x}) = \max_{i=1}^n x_i$ je maximum ze složek \mathbf{x} . Konvexitu této funkce lze dokázat z Definice 11.2, nicméně dokažme ji z Věty 11.7. Máme $g_i(\mathbf{x}) = x_i$. Funkce g_i jsou lineární, tedy konvexní. Tedy funkce $f(\mathbf{x}) = \max_{i=1}^n g_i(\mathbf{x})$ je konvexní. \square

Dále uveďme příklady pro nekonečnou množinu I .

Příklad 11.7. 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 11.6 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} – pro zdůraznění této skutečnosti můžeme psát $\|\mathbf{x} - \mathbf{y}\| = g_{\mathbf{y}}(\mathbf{x})$. Jelikož f je maximum těchto funkcí, je i funkce f konvexní. \square

Příklad 11.8. 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. \square

Příklad 11.9. 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 §5.8) znamená vypočítat

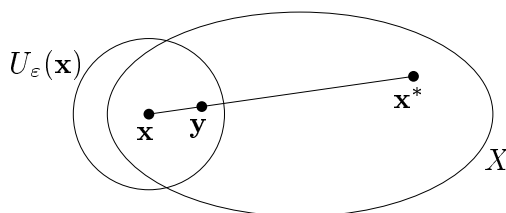
$$f(\mathbf{w}) = \min_{\mathbf{x} \in \mathbb{R}^m} \sum_{i=1}^n w_i (\mathbf{a}_i^T \mathbf{x} - b_i)^2,$$

kde jsme označili hodnotu výsledného minima jako funkci vektoru vah. Funkce f je konkávní, protože je minimem lineárních funkcí. \square

11.3 Konvexní optimalizační úlohy

Věta 11.8. Necht' funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$. Pak každé lokální minimum funkce f na množině X je zároveň globální.

Důkaz. Necht' \mathbf{x} je lokálním minimem f na X , viz obrázek:



Dle Definice 9.2 tedy existuje $\varepsilon > 0$ tak, že $f(\mathbf{x}) \leq f(\mathbf{y})$ pro všechna $\mathbf{y} \in U_\varepsilon(\mathbf{x}) \cap X$. Necht' ale \mathbf{x} není globální minimum, tedy existuje $\mathbf{x}^* \in X$ takové, že $f(\mathbf{x}^*) < f(\mathbf{x})$. Ukážeme, že to vede ke sporu. Pro každé ε totiž můžeme zvolit $0 < \alpha < 1$ tak, že bod $\mathbf{y} = \alpha\mathbf{x} + (1 - \alpha)\mathbf{x}^*$ leží v okolí $U_\varepsilon(\mathbf{x})$. Protože je množina X konvexní, leží bod \mathbf{y} zároveň i v X . Máme

$$f(\mathbf{y}) = f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{x}^*) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{x}^*) < \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{x}) = f(\mathbf{x}).$$

Ale tvrzení $f(\mathbf{y}) < f(\mathbf{x})$ je ve sporu s předpokladem, že \mathbf{x} je lokální minimum. \square

Zopakujme nyní obecnou úlohu spojitě optimalizace (1.4)

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmíněk} \quad & g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \\ & h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, l \end{aligned} \tag{11.7}$$

kde $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $(g_1, \dots, g_m) = \mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $(h_1, \dots, h_l) = \mathbf{h}: \mathbb{R}^n \rightarrow \mathbb{R}^l$.

Definice 11.3. *Konvexní optimalizační úloha je úloha (11.7), kde funkce f, g_1, \dots, g_m jsou konvexní a funkce h_1, \dots, h_l jsou afinní (tedy zobrazení \mathbf{h} je afinní).*

Zde je nejasnost: má být f konvexní na celém definičním oboru, nebo jen na množině přípustných řešení?

Množina přípustných řešení konvexní úlohy je konvexní. Můžeme ji totiž psát jako

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \} = \bigcap_{i=1}^m \{ \mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0 \} \cap \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{h}(\mathbf{x}) = \mathbf{0} \}.$$

Zde každá množina $\{ \mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0 \}$ je konvexní, neboť je to subkontura konvexní funkce g_i (Věta 11.2). Množina $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{h}(\mathbf{x}) = \mathbf{0} \}$ je afinní podprostor, tedy také konvexní. Množina X je průnik konvexních množin, tedy (dle Věty 11.1) je konvexní.

Mohli bychom si myslet, že je přirozenější definovat konvexní optimalizační úlohu jednoduše jako minimalizaci konvexní funkce na konvexní množině. Tato definice je obecnější, protože množina přípustných řešení X může být konvexní i tehdy, když funkce g_i nejsou konvexní nebo funkce h_i nejsou afinní. Výhoda Definice 11.3 je v tom, že zatímco konvexitu množiny X nemusí být snadné dokázat, obvykle je snadno vidět, zda jsou funkce g_i konvexní a funkce h_i afinní.

Příklad 11.10. Uvažujme dvě ekvivalentní definice téže množiny

$$X = \{ \mathbf{x} \in \mathbb{R}^2 \mid x_1/(1 + x_2^2) \leq 0, (x_1 + x_2)^2 = 0 \} = \{ \mathbf{x} \in \mathbb{R}^2 \mid x_1 \leq 0, x_1 + x_2 = 0 \}.$$

Oba tvary jsou ekvivalentní (proč?). V prvním tvaru funkce $g(\mathbf{x}) = x_1/(1 + x_2^2)$ není konvexní (dokažte z Definice 11.2!) a funkce $h(\mathbf{x}) = (x_1 + x_2)^2$ není afinní. Přesto je množina X konvexní, což je vidět ze druhého tvaru. \square

Je obvykle poměrně snadné najít (nějaké) lokální minimum optimalizační úlohy (at' konvexní či nekonvexní). Pro úlohy bez omezení s diferencovatelnou účelovou funkcí to lze udělat např. gradientní metodou (viz §10.4). Numerické algoritmy pro úlohy s omezeními existují, ale neuvádějí je. Konvexní optimalizační úlohy se těší výsadě dané Větou 11.8, totiž že každé lokální minimum je zároveň globální. Pokud je úloha nekonvexní, obvykle (avšak ne vždy, viz Příklad 15.5) má mnoho lokálních minim a kvůli tomu je těžké najít globální optimum. Nekonvexní úloha s větším množstvím proměnných je tedy velmi často prakticky neřešitelná.

11.4 Cvičení

11.1. Dokažte z definice konvexní množiny, že následující množiny jsou konvexní:

- a) interval $[a, b] \subseteq \mathbb{R}$
- b) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \leq \mathbf{b}, \mathbf{Cx} = \mathbf{d} \}$
- c) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{Ax} \leq 1 \}$, kde \mathbf{A} je pozitivně semidefinitní

11.2. Které z následujících množin jsou konvexní? Nemusíte dokazovat z definice, stačí uvést přesvědčivý argument. Pokud to jde, zkuste množinu načrtnout v prostoru malé dimenze.

- a) $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1 \}$ (nadrovina, konvexní)
- b) $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i \geq 1 \}$ (poloprostor, konvexní)
- c) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_{i=1}^n x_i = 1 \}$ (průnik poloprostorů a nadroviny, tedy konvexní polyedr)
- d) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_{i=1}^n x_i \leq 1 \}$ (průnik poloprostorů, konvexní)
- e) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 = 1 \}$ (sféra, není konvexní)
- f) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 < 1 \}$ (koule bez hranice, konvexní)
- g) $\{ (x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, xy = 1 \}$ (graf jedné větve hyperboly, není konvexní)
- h) $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 2 \} \cap \{ (x, y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 \leq 2 \}$ (průnik dvou koulí, konvexní)

11.3. Pro každou funkci dokažte z Definice 11.2, které z těchto čtyřech tvrzení platí: funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní.

- a) $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
- b) $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$
- c) $f(\mathbf{x}) = \max_{i=1}^n x_i$
- d) $f(\mathbf{x}) =$ aritmetický průměr čísel x_1, \dots, x_n

11.4. Pro každou funkci dokažte, které z těchto čtyřech tvrzení platí: funkce je konvexní, konkávní, konvexní i konkávní, ani konvexní ani konkávní. Můžete to dokázat buď z Definice 11.2, pomocí derivací, nebo pomocí operací zachovávajících konvexitu.

- a) $f(x) = e^{x^2}$
- b) $f(x) = e^{-x^2}$
- c) $f(x, y) = |x - y|$
- d) $f(x, y) = -y$
- e) $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|^2$
- f) $f(\mathbf{x}) = \sum_{i=1}^n x_i \log x_i$ na množině \mathbb{R}_{++}^n
- g) $f(\mathbf{x}) = \sum_{i=1}^k \log(b_i - \mathbf{x}^T \mathbf{a}_i)$ na množině $X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{a}_i < b_i, i = 1, \dots, k \}$
- h) $f(\mathbf{x}) = \max_{i=1}^n |x_i|$
- i) $f(\mathbf{x}) = \min_{i=1}^n |x_i|$
- j) $f(\mathbf{x}) = \max_{i=1}^n x_i + \min_{i=1}^n x_i$
- k) $f(\mathbf{x}) = \max_{i=1}^n x_i - \min_{i=1}^n x_i$

- l) (\star) $f(\mathbf{x}) = \text{median}_{i=1}^n x_i$ (medián čísel x_1, \dots, x_n)
 m) (\star) $f(\mathbf{x}) = \text{součet } k \text{ největších čísel } x_1, \dots, x_n$ (kde $k \leq n$ je dáno)

11.5. Robustní prokládání přímky množinou bodů $(\mathbf{x}_i, y_i) \in (\mathbb{R}^n \times \mathbb{R})$, $i = 1, \dots, m$ vyžaduje minimalizaci funkce

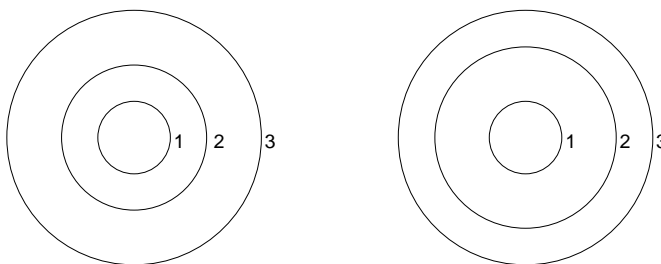
$$f(\mathbf{a}, b) = \sum_{i=1}^m \max\{-\mathbf{a}^T \mathbf{x}_i + b + y_i - \varepsilon, 0, \mathbf{a}^T \mathbf{x}_i + b - y_i - \varepsilon\},$$

kde $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Dokažte, že $f(\mathbf{a}, b)$ je konvexní funkce.

11.6. Je dána funkce $f(x) = -\cos x$ a množina $X = [-\pi, +\pi]$ (kde $[\cdot]$ značí uzavřený interval). Zakroužkujte pravdivá tvrzení (může jich být i více):

- a) Funkce f je na množině X konvexní.
 b) Funkce f je na množině X konkávní.
 c) Funkce f není na množině X ani konvexní ani konkávní.

11.7. Každý z obrázků zobrazuje některé vrstevnice funkce dvou proměnných a jejich výšky. Je možné, aby funkce, která má tyto vrstevnice, byla konvexní? Dokažte z Definice 11.2. (Odpověď: ne, ano)



11.8. Významnou vlastností konvexních funkcí je to, že každé lokální minimum funkce je zároveň globální (Věta 11.8). Ne každá funkce s touto vlastností je ovšem konvexní. Člověk by si mohl myslet, že součet dvou funkcí (ne nutně konvexních) s touto vlastností bude mít tuto vlastnost také. Je toto tvrzení pravdivé? Odpověď dokažte.

11.9. Dokažte, že množina optimálních řešení konvexní optimalizační úlohy je konvexní.

11.10. 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?

11.11. Bude Věta 11.1 platit, pokud v ní sousloví ‘konvexní množina’ nahradíme souslovím ‘lineární podprostor’ (příp. ‘afinní podprostor’, ‘konvexní kužel’)? Kladnou i zápornou odpověď dokažte.

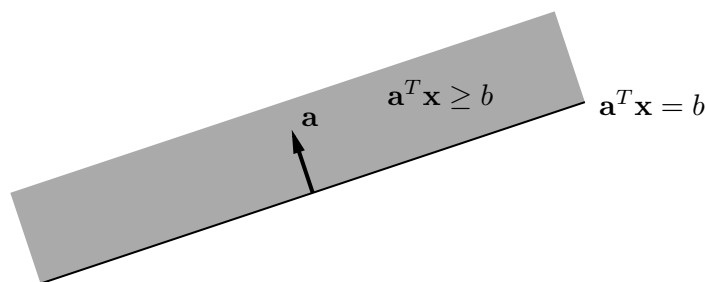
Kapitola 12

Lineární programování

12.1 Konvexní polyedry

Motivace: Jak byste řešili soustavu lineárních nerovnic? Jaká je jejich množina řešení?

Poloprostor je množina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \geq b\}$ pro nějaké $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$. Jeho hranice je nadrovina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b\}$. Obrázek znázorňuje tyto pojmy pro $n = 2$:



Definice 12.1. Konvexní polyedr je průnik konečně mnoha poloprostorů.

Konvexní polyedr je tedy množina

$$X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} \geq b_i, i = 1, \dots, m\} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}, \quad (12.1)$$

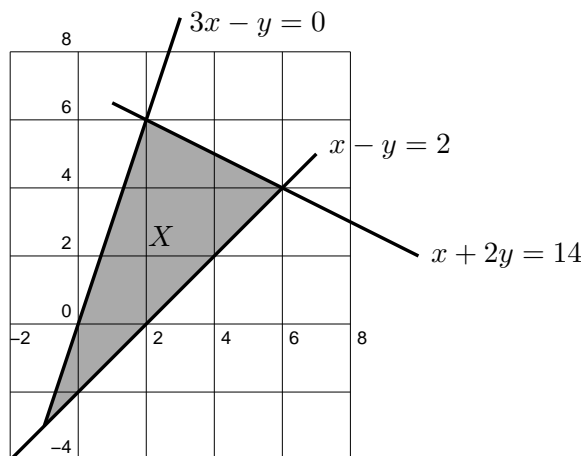
kde $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ jsou řádky matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $b_1, \dots, b_m \in \mathbb{R}$ jsou složky vektoru $\mathbf{b} \in \mathbb{R}^m$. Je jasné, že definice dovoluje i omezení typu $\mathbf{a}_i^T \mathbf{x} \leq b_i$, protože to je ekvivalentní $-\mathbf{a}_i^T \mathbf{x} \geq -b_i$. Dovoluje i omezení typu rovnosti $\mathbf{a}_i^T \mathbf{x} = b_i$, které je ekvivalentní $\mathbf{a}_i^T \mathbf{x} \leq b_i$, $\mathbf{a}_i^T \mathbf{x} \geq b_i$.

Jelikož poloprostor je očividně konvexní množina, plyne konvexita konvexního polyedru z Věty 11.1. Všimněte si, že konvexní polyedr nemusí být omezený.

Příklad 12.1. Množina

$$X = \{(x, y) \in \mathbb{R}^2 \mid x + 2y \leq 14, 3x - y \geq 0, x - y \leq 2\}$$

je konvexní polyedr, který snadno nakreslíme:



□

Příklad 12.2. Příklady konvexních polyedrů v \mathbb{R}^n :

- prázdná množina \emptyset
- celý prostor \mathbb{R}^n
- každý afinní podprostor (např. bod, přímka, rovina, nadrovina)
- polopřímka $\{ \mathbf{x} + \alpha \mathbf{v} \mid \alpha \geq 0 \}$
- poloprostor
- panel $\{ \mathbf{x} \in \mathbb{R}^n \mid b_1 \leq \mathbf{a}^T \mathbf{x} \leq b_2 \}$
- hyperkrychle $\{ \mathbf{x} \in \mathbb{R}^n \mid -1 \leq x_i \leq 1, i = 1, \dots, n \}$
- simplex, to jest konvexní obal $n + 1$ afinně nezávislých bodů
- standardní simplex $\{ \mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i \leq 1 \}$
- pravděpodobnostní simplex $\{ \mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = 1 \}$ (množina všech rozdělení pravděpodobnosti diskrétní náhodné proměnné)
- zobecněný osmistěn $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n |x_i| \leq 1 \}$.

□

12.1.1 Stěny konvexního polyedru

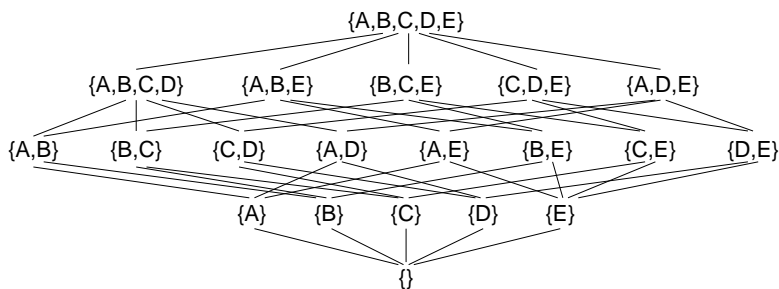
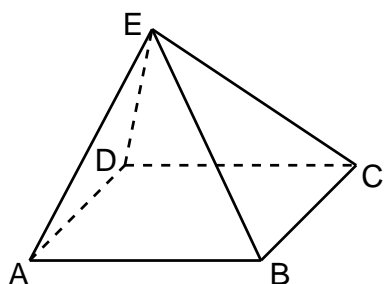
Definice 12.2. Necht' $I \subseteq \{1, \dots, m\}$. **Stěna konvexního polyedru (12.1)** je množina všech jeho bodů, které splňují $\mathbf{a}_i^T \mathbf{x} = b_i$ pro všechna $i \in I$ (tedy ve kterých jsou podmínky I aktivní). Prázdná množina je také stěna.

Volbou $I = \emptyset$ dostaneme, že jednou ze stěn je i celý polyedr. **Dimenze stěny** je dimenze jejího afinního obalu (zopakujte si pojem afinního obalu z §11.1.1 a dimenze afinního podprostoru z §3.3). Stěny některých dimenzí mají jméno:

- stěna dimenze 0 se nazývá **vrchol**,
- stěna dimenze 1 se nazývá **hrana**,
- stěna dimenze $n - 1$ se nazývá **faceta** (angl. *facet*, zatímco *face* znamená stěna).

Z Definice 12.2 snadno plyne, že každá stěna konvexního polyedru je sama o sobě konvexní polyedr. Stěny konvexního polyedru tvoří částečně uspořádanou množinu indukovanou množinovou inkluzí (tzv. *svaz stěn*).

Příklad 12.3. Konvexní polyedr v \mathbb{R}^3 a jeho svaz stěn.



□

12.1.2 (*) Dvě reprezentace polyedru

Následující věta je hluboká a uvádíme ji bez důkazu. (Pro neomezené konvexní polyedry platí podobná věta, trochu složitější, kterou neuvádíme.)

Věta 12.1. *Konvexní obal konečně mnoha bodů je omezený konvexní polyedr. Obráceně, omezený konvexní polyedr je konvexním obalem svých vrcholů.*

Máme tedy dvě reprezentace omezeného polyedru:

- **H-reprezentace:** průnik konečně mnoha poloprostorů ('H' jako *half-space*)
- **V-reprezentace:** konvexní obal konečně mnoha bodů ('V' jako *vertex*)

Přechod od jedné reprezentace ke druhé může být výpočetně těžký či prakticky nemožný. Důvodem je to, že polyedr definovaný jako průnik malého počtu (přesněji, tento počet je polynomiální funkcí n) poloprostorů může mít velmi velký (exponenciální funkce n) počet vrcholů. Naopak, polyedr s malým počtem vrcholů může mít exponenciální počet facet. V tom případě by algoritmus, který převádí H -reprezentaci na V -reprezentaci nebo naopak, by při polynomiálně dlouhém vstupu musel vydat exponenciálně dlouhý výstup.

Příklad 12.4.

- Simplex (tedy konvexní obal $n + 1$ bodů) je konvexní polyedr, který má $n + 1$ vrcholů a $n + 1$ facet.
- Hyperkrychle má $2n$ facet a 2^n vrcholů.
- Zobecněný osmistěn $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_i |x_i| \leq 1 \}$ má $2n$ vrcholů a 2^n facet. □

12.2 Úloha lineárního programování

Úloha **lineárního programování** (LP, také zvané lineární optimalizace) znamená minimalizaci lineární funkce za podmínek lineárních rovností a nerovností. Přesněji, v obecné formulaci (1.4) je funkce f lineární a funkce g_i, h_i jsou afinní.

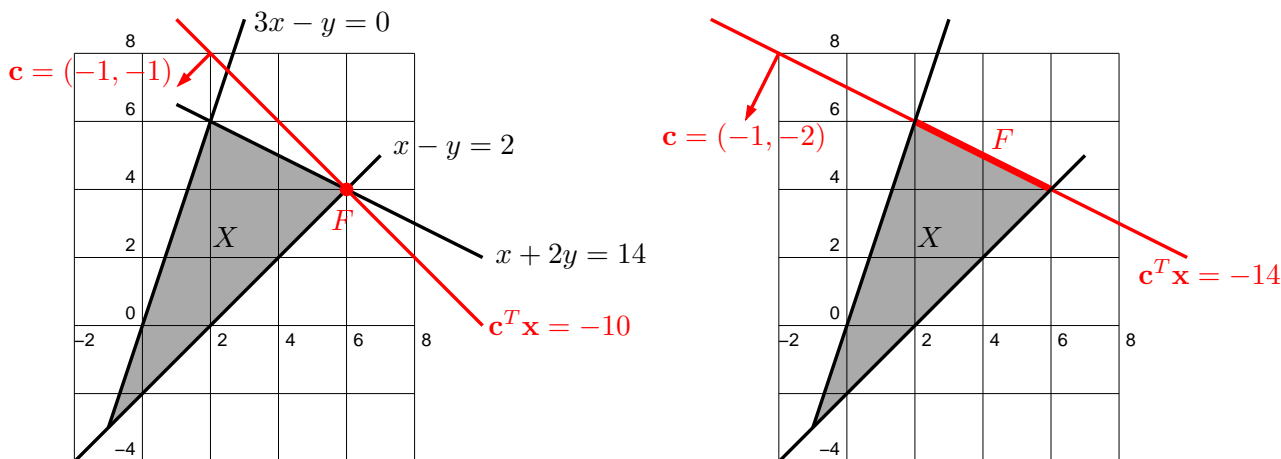
Z definice lineárního programování plyne, že jeho množinou přípustných řešení je konvexní polyedr. Zopakujme (viz §1.2), že pro řešitelnost LP mohou nastat tři případy: úloha má (alespoň jedno) optimální řešení, úloha je *nepřípustná* (tj. množina přípustných řešení je prázdná, neboli podmínky si navzájem odporují), úloha je *neomezená* (tj. účelovou funkci lze za splněných podmínek zlepšovat nade všechny meze).

Jednoduché úlohy lineárního programování lze řešit graficky.

Příklad 12.5. Mějme lineární program

$$\begin{aligned} \min \quad & -x - y \\ \text{za podmíněk} \quad & x + 2y \leq 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \end{aligned} \tag{12.2}$$

Množina přípustných řešení X této úlohy je konvexní polyedr z Příkladu 12.1:



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 vektoru \mathbf{c} . Je proto jasné (viz levý obrázek), že účelová funkce na množině X nabývá (globálního) minima v bodě $(x, y) = (6, 4)$. Tento bod je vrchol polyedru (stěna dimenze 1). Úloha má právě jedno optimální řešení.

Pokud bychom účelovou funkci úlohy (12.2) změnili na $-x - 2y$, bude tato funkce na množině X nabývat minima ve všech bodech úsečky spojující body $(2, 6)$ a $(6, 4)$ (viz pravý obrázek), tedy na stěně polyedru dimenze 2. Úloha má nekonečně mnoho optimálních řešení. \square

Věta 12.2. *Minimum lineární funkce na konvexním polyedru se nabývá na stěně tohoto polyedru. Neboli, je-li $X \subseteq \mathbb{R}^n$ konvexní polyedr a $\mathbf{c} \in \mathbb{R}^n$, pak množina $\operatorname{argmin}_{\mathbf{x} \in X} \mathbf{c}^T \mathbf{x}$ je stěna polyedru X .*

12.3 Různé tvary úloh LP

Při zápisu úlohy LP je zvykem odděleně zapisovat obecná lineární omezení a omezení na znaménka jednotlivých proměnných. Obecnou úlohu LP tedy zapíšeme jako

$$\begin{aligned} \min \quad & c_1 x_1 + \cdots + c_n x_n \\ \text{za podmíněk} \quad & a_{i1} x_1 + \cdots + a_{in} x_n \geq b_i, \quad i \in I_+ \\ & a_{i1} x_1 + \cdots + a_{in} x_n \leq b_i, \quad i \in I_- \\ & a_{i1} x_1 + \cdots + a_{in} x_n = b_i, \quad i \in I_0 \\ & x_j \geq 0, \quad j \in J_+ \\ & x_j \leq 0, \quad j \in J_- \\ & x_j \in \mathbb{R}, \quad j \in J_0 \end{aligned}$$

kde

$$I = \{1, \dots, m\} = I_0 \cup I_+ \cup I_-$$

$$J = \{1, \dots, n\} = J_0 \cup J_+ \cup J_-$$

jsou rozklady indexových množin. Zápis $x_j \geq 0$ značí, že proměnná x_j může nabývat pouze nezáporných hodnot, zatímco $x_j \in \mathbb{R}$ značí, že x_j může nabývat libovolných hodnot.

Počítačové algoritmy na řešení LP často předpokládají úlohu v nějakém speciálním tvaru, kdy jsou dovoleny pouze jisté typy omezení. Nejčastěji užívané speciální tvary jsou:

- Dovolíme pouze omezení typu ‘=’ a nezáporné proměnné ($I_+ = I_- = J_- = J_0 = \emptyset$), tj.

$$\begin{aligned} \min \quad & c_1 x_1 + \dots + c_n x_n \\ \text{za podmíněk} \quad & a_{i1} x_1 + \dots + a_{in} x_n = b_i, \quad i = 1, \dots, m \\ & x_j \geq 0, \quad j = 1, \dots, n \end{aligned}$$

To¹ lze psát maticově jako $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, kde $\mathbf{x}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$.

- Tvar $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.
- Tvar $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$.

Tyto speciální tvary nemají menší vyjadřovací schopnost než obecný tvar, neboť obecný tvar se dá převést na libovolný speciální tvar některou z následujících úprav:

- Rovnost $\mathbf{a}_i^T \mathbf{x} = b_i$ nahradíme dvěma nerovnostmi $\mathbf{a}_i^T \mathbf{x} \geq b_i$, $-\mathbf{a}_i^T \mathbf{x} \geq -b_i$.
- Nerovnost $a_{i1}x_1 + \dots + a_{in}x_n \leq b_i$ převedeme na rovnost přidáním pomocné **slackové proměnné**² $u_i \geq 0$ jako $a_{i1}x_1 + \dots + a_{in}x_n + u_i = b_i$.

Podobně převedeme nerovnost $a_{i1}x_1 + \dots + a_{in}x_n \geq b_i$ na rovnost (jak?).

- Proměnnou bez omezení $x_i \in \mathbb{R}$ rozdělíme na dvě nezáporné proměnné $x_i^+ \geq 0$, $x_i^- \geq 0$ přidáním podmínky $x_i = x_i^+ - x_i^-$.

Úloha získaná z původní úlohy pomocí těchto úprav je ekvivalentní původní úloze v tom smyslu, že hodnota jejich optima je stejná a argument optima původní úlohy lze ‘snadno’ získat z argumentu optima nové úlohy.

Příklad 12.6. V úloze (12.2) chceme první podmínku převést na rovnost. To uděláme zavedením slackové proměnné $u \geq 0$. Transformovaná úloha je

$$\begin{aligned} \min \quad & -x - y \\ \text{za podmíněk} \quad & x + 2y + u = 14 \\ & 3x - y \geq 0 \\ & x - y \leq 2 \\ & u \geq 0 \end{aligned}$$

Je-li (x, y, u) optimum této úlohy, optimum úlohy (12.2) je (x, y) . □

¹ Tomuto tvaru se někdy říká *standardní*. Bohužel názvosloví různých tvarů LP není jednotné, názvy jako ‘standardní tvar’, ‘základní tvar’ či ‘kanonický tvar’ tedy mohou znamenat v různých knihách něco jiného.

² *Slack* 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*.

Příklad 12.7. V úloze (12.2) obě proměnné mohou mít libovolné znaménko. Chceme převést úlohu na tvar, kde všechny proměnné jsou nezáporné. Dosadíme $x = x_+ - x_-$ a $y = y_+ - y_-$, kde $x_+, x_-, y_+, y_- \geq 0$. Výsledná úloha je

$$\begin{aligned} \min \quad & -x_+ + x_- - y_+ + y_- \\ \text{za podmínek} \quad & x_+ - x_- + 2y_+ - 2y_- \leq 14 \\ & 3x_+ - 3x_- - y_+ + y_- \geq 0 \\ & x_+ - x_- - y_+ + y_- \leq 2 \\ & x_+, x_-, y_+, y_- \geq 0 \end{aligned} \quad \square$$

12.3.1 Po částech afinní funkce

Mějme účelovou funkci

$$f(\mathbf{x}) = \max_{k=1}^K (\mathbf{c}_k^T \mathbf{x} + d_k), \quad (12.3)$$

kde $\mathbf{c}_k \in \mathbb{R}^n$ a $d_k \in \mathbb{R}$ jsou dány (viz Cvičení 12.7). Řešme úlohu

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{Ax} \geq \mathbf{b} \end{aligned} \quad (12.4)$$

Toto není úloha LP, neboť funkce f není lineární nebo afinní, je pouze po částech afinní. Ovšem podle Věty 11.7 je tato funkce *konvexní*, proto úloha (12.4) je minimalizace konvexní funkce na konvexním polyedru a tedy se na ní vztahuje Věta 11.8.

Úloha jde převést na LP zavedením pomocné proměnné:

$$\begin{aligned} \min \quad & z \\ \text{za podmínek} \quad & \mathbf{c}_k^T \mathbf{x} + d_k \leq z, \quad k = 1, \dots, K \\ & \mathbf{Ax} \geq \mathbf{b} \end{aligned} \quad (12.5)$$

kde minimalizujeme přes proměnné $(x_1, \dots, x_n, z) = (\mathbf{x}, z) \in \mathbb{R}^{n+1}$. Ekvivalence úloh (12.4) a (12.5) se dokáže takto. Předpokládejme, že (\mathbf{x}, z) je optimum úlohy (12.5). Pak musí být alespoň jedno z omezení $\mathbf{c}_k^T \mathbf{x} + d_k \leq z$ aktivní (tedy musí platit s rovností), protože jinak bychom mohli z zmenšit a neporušit přitom žádné omezení. Z toho plyne $z = \max_{k=1}^K (\mathbf{c}_k^T \mathbf{x} + d_k)$.

Tento převod lze užít i na funkce obsahující absolutní hodnoty, neboť $|x| = \max\{-x, x\}$.

Při těchto převodech je nutná opatrnost: pokud bychom v úloze (12.4) maximalizovali místo minimalizovali, převod na LP by nebyl možný. (Vodítkem je, že pokud úloha nejde jednoduše převést na minimalizaci konvexní funkce na konvexním polyedru, nepůjde převést na LP.)

Příklad 12.8. Úloha

$$\begin{aligned} \min \quad & \max\{3x_1 + 4x_2, 2x_1 - 3x_2\} \\ \text{za podm.} \quad & x_1 + 2x_2 \leq 14 \\ & 3x_1 - x_2 \geq 0 \\ & x_1 - x_2 \leq 2 \end{aligned}$$

není LP, protože účelová funkce $f(x_1, x_2) = \max\{3x_1 + 4x_2, 2x_1 - 3x_2\}$ není lineární ani afinní. Lze ji ale přeformulovat na

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & 3x_1 + 4x_2 \leq z \\ & 2x_1 - 3x_2 \leq z \\ & x_1 + 2x_2 \leq 14 \\ & 3x_1 - x_2 \geq 0 \\ & x_1 - x_2 \leq 2 \end{aligned}$$

což je LP, neboť účelová funkce $f(x_1, x_2, z) = z$ je lineární a omezení jsou také lineární. \square

Podobně lze často na LP převést úlohy, které obsahují minima a maxima v omezeních.

Příklad 12.9. Platí rovnost

$$\min\{x - y \mid x \geq 0, y \geq 0, \max\{x, y\} \leq 1\} = \min\{x - y \mid x \geq 0, y \geq 0, x \leq 1, y \leq 1\}$$

protože $\max\{x, y\} \leq 1$ je ekvivalentní $x \leq 1, y \leq 1$. Úloha vlevo není LP, úloha vpravo ano. \square

12.4 Některé aplikace LP

12.4.1 Optimální výrobní program

Z m druhů surovin vyrábíme n druhů výrobků.

- a_{ij} = množství suroviny druhu i potřebné na výrobu výrobku druhu j
- b_i = množství suroviny druhu i , které máme k dispozici
- c_j = zisk z vyrobení jednoho výrobku druhu j
- x_j = počet vyrobených výrobků druhu j

Úkolem je zjistit, kolik jakých výrobků máme vyrobit, abychom dosáhli největšího zisku. Řešení:

$$\max \left\{ \sum_{j=1}^n c_j x_j \mid \sum_{j=1}^n a_{ij} x_j \leq b_i, x_j \geq 0 \right\}. \quad (12.6)$$

Příklad 12.10. Pán u stánku prodává lupínky za 120 Kč/kg a hranolky za 76 Kč/kg. Na výrobu 1 kg lupínků se spotřebuje 2 kg brambor a 0.4 kg oleje. Na výrobu 1 kg hranolku se spotřebuje 1.5 kg brambor a 0.2 kg oleje. Je nakoupeno 100 kg brambor a 16 kg oleje. Brambory stály 12 Kč/kg, olej 40 Kč/kg. Kolik má pán vyrobit lupínků a kolik hranolků, aby co nejvíce vydělal? To lze vyjádřit jako LP

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

Přitom předpokládáme, že zbytky surovin se po pracovní době vyhodí. Pokud se zbytky využijí, tak maximalizujeme $(120 - 24 - 16)l + (76 - 18 - 8)h = 80l + 50h$.

V obou případech je optimální řešení $l = 20$ kg lupínků a $h = 40$ kg hranolků. \square

12.4.2 Směšovací (dietní) problém

Z n druhů surovin, z nichž každá je směsí m druhů látek, máme namíchat konečný produkt o požadovaném složení tak, aby cena surovin byla minimální.

- a_{ij} = množství látky druhu i obsažené v jednotkovém množství suroviny druhu j
- b_i = nejmenší požadované množství látky druhu i v konečném produktu
- c_j = jednotková cena suroviny druhu j
- x_j = množství suroviny druhu j

Řešení:

$$\min \left\{ \sum_{j=1}^n c_j x_j \mid \sum_{j=1}^n a_{ij} x_j \geq b_i, x_j \geq 0 \right\}. \quad (12.7)$$

Příklad 12.11. 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?

Minimalizujeme $25b + 50m + 80z$ za podmínek $2b + 1m + 1z \geq 8$, $2b + 6m + 1z \geq 16$, $1b + 3m + 6z \geq 8$ a $b, m, z \geq 0$. Optimální řešení je $b = 3.2$, $m = 1.6$, $z = 0$ s hodnotou 160. \square

12.4.3 Dopravní problém

Máme m výrobců a n spotřebitelů.

- a_i = množství zboží vyráběné výrobcem i
- b_j = množství zboží požadované spotřebitelem j
- c_{ij} = cena dopravy jednotky zboží od výrobce i ke spotřebiteli j
- x_{ij} = množství zboží vezené od výrobce i ke spotřebiteli j

Chceme co nejlevněji rozvézt zboží od výrobců ke spotřebitelům. Řešení:

$$\min \left\{ \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \mid \sum_{j=1}^n x_{ij} = a_i, \sum_{i=1}^m x_{ij} = b_j, x_{ij} \geq 0 \right\}. \quad (12.8)$$

Zadání musí splňovat $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$ (nabídka musí být rovna poptávce), jinak bude úloha nepřijatelná. Úloha jde modifikovat tak, že dovolíme $\sum_{i=1}^m a_i \geq \sum_{j=1}^n b_j$ (proved'te!).

12.4.4 Distribuční problém

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:

$$\min \left\{ \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \mid \sum_{j=1}^n x_{ij} \leq a_i, \sum_{i=1}^m k_{ij} x_{ij} = b_j, x_{ij} \geq 0 \right\}. \quad (12.9)$$

Sem přidat příklad aplikace LP na grafové úlohy: nejkratší cesta, maximální tok

12.5 Přeuročené lineární soustavy

12.5.1 Vektorové normy

Norma formalizuje pojem ‘délky’ vektoru \mathbf{x} .

Definice 12.3. Funkce $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ se nazývá vektorová³ norma, jestliže splňuje tyto axiomy:

1. Jestliže $\|\mathbf{x}\| = 0$ pak $\mathbf{x} = \mathbf{0}$.
2. $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ pro každé $\alpha \in \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$ (norma je kladně homogenní).
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (trojúhelníková nerovnost).

Z axiomů plynou tyto další vlastnosti normy:

- $\|\mathbf{0}\| = 0$, což plyne z homogenity pro $\alpha = 0$
- $\|\mathbf{x}\| \geq 0$ pro každé $\mathbf{x} \in \mathbb{R}^n$. To jde odvodit tak, že v trojúhelníkové nerovnosti položíme $\mathbf{y} = -\mathbf{x}$, což dá

$$\|\mathbf{x} - \mathbf{x}\| = \|\mathbf{0}\| = 0 \leq \|\mathbf{x}\| + \|-\mathbf{x}\| = 2\|\mathbf{x}\|,$$

kde na pravé straně jsme použili homogenitu.

- Norma je konvexní funkce, neboť pro každé $0 \leq \alpha \leq 1$ máme

$$\|\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}\| \leq \|\alpha\mathbf{x}\| + \|(1 - \alpha)\mathbf{y}\| = \alpha\|\mathbf{x}\| + (1 - \alpha)\|\mathbf{y}\|,$$

kde nerovnost plyne z trojúhelníkové nerovnosti a rovnost z homogenity.

Jednotková sféra normy je množina $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$, tedy vrstevnice normy jednotkové výšky. Díky homogenitě je jednotková sféra středově symetrická a její tvar zcela určuje normu.

Uveďme příklady norem. Základním příkladem je **p -norma**

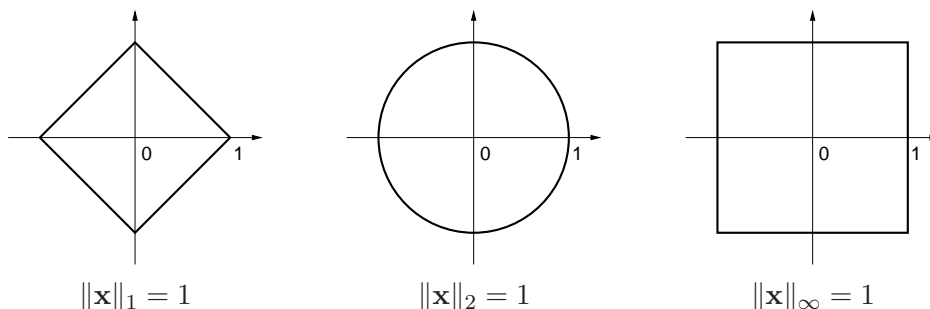
$$\|\mathbf{x}\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p}.$$

Musí být $p \geq 1$, jinak neplatí trojúhelníková nerovnost. Nejčastěji narazíte na:

³ Existují i *maticové* normy, což jsou funkce $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}$.

- $\|\mathbf{x}\|_1 = |x_1| + \dots + |x_n|$. 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 známá *eukleidovská norma*.
- $\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \max\{|x_1|, \dots, |x_n|\}$ (dokažte rovnost výpočtem limity!). Někdy se jí říká *Čebyševova norma* nebo *max-norma*.

Jednotkové sféry těchto norem v \mathbb{R}^2 vypadají takto:



Existují ale i normy, které nejsou p -normy, např.

- $\|\mathbf{x}\| = 2|x_1| + \sqrt{x_2^2 + x_3^2} + \max\{|x_4|, |x_5|\}$ je norma na \mathbb{R}^5 .
- Je-li $\|\mathbf{x}\|$ norma a \mathbf{A} je čtvercová nebo úzká matice s plnou hodnotí, je také $\|\mathbf{Ax}\|$ norma.

12.5.2 Přibližné řešení přeurených lineárních soustav

Mějme přeurenou lineární soustavu $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ a $\mathbf{0} \neq \mathbf{b} \in \mathbb{R}^m$. Nalezení jejího přibližného řešení formulujeme jako úlohu

$$\min\{\|\mathbf{Ax} - \mathbf{b}\|_p \mid \mathbf{x} \in \mathbb{R}^n\}. \quad (12.10)$$

Uvažujme tři případy:

- Pro $p = \infty$ hledáme takové \mathbf{x} , které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_\infty = \max_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (12.11)$$

tedy minimalizuje maximální residuum. Toto řešení je známé pod názvem *minimaxní* nebo *Čebyševovo*. Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & \mathbf{a}_i^T \mathbf{x} - b_i \leq z, \quad i = 1, \dots, m \\ & -\mathbf{a}_i^T \mathbf{x} + b_i \leq z, \quad i = 1, \dots, m \end{aligned}$$

který lze zapsat elegantněji jako

$$\min\{z \in \mathbb{R} \mid \mathbf{x} \in \mathbb{R}^n, -z\mathbf{1} \leq \mathbf{Ax} - \mathbf{b} \leq z\mathbf{1}\}. \quad (12.12)$$

- 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

$$\|\mathbf{Ax} - \mathbf{b}\|_1 = \sum_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|, \quad (12.13)$$

kde $\mathbf{a}_1, \dots, \mathbf{a}_m$ jsou řádky matice \mathbf{A} . Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & \sum_{i=1}^m z_i \\ \text{za podm.} \quad & \mathbf{a}_i^T \mathbf{x} - b_i \leq z_i, \quad i = 1, \dots, m \\ & -\mathbf{a}_i^T \mathbf{x} + b_i \leq z_i, \quad i = 1, \dots, m \end{aligned}$$

který lze zapsat elegantněji v maticovém tvaru jako

$$\min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{z} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^n, -\mathbf{z} \leq \mathbf{Ax} - \mathbf{b} \leq \mathbf{z}\}. \quad (12.14)$$

12.5.3 Použití na robustní regresi

Řešení ve smyslu 1-normy se používá tehdy, když potřebujeme modelovat funkční závislost naměřených dat (tedy děláme regresi, viz §5.3.1) a malá část dat je naměřená úplně špatně (např. se někdo při zapisování čísel spletl v desetinné čárce). Takovým datovým bodům s hrubou chybou se říká **vychýlené body** (*outliers*). Disciplína zabývající se modelováním funkčních závislostí za přítomnosti vychýlených bodů se nazývá **robustní regrese**.

V tomto případě řešení ve smyslu nejmenších čtverců není vhodné (není 'robustní'), protože i jediný vychýlený bod velmi ovlivní řešení. Řešení ve smyslu 1-normy tuto neblahou vlastnost nemá, přesněji, má ji v menší míře.

Ukážeme to na nejjednodušším možném případě regrese: odhad hodnoty jediného reálného čísla ze souboru jeho nepřesných měření. Mějme čísla $\mathbf{b} = (b_1, \dots, b_m) \in \mathbb{R}^m$ a řešme úlohu (12.10) ve tvaru

$$\min\{\|(x - b_1, \dots, x - b_m)\|_p \mid x \in \mathbb{R}\} = \min\{\|\mathbf{1}x - \mathbf{b}\|_p \mid x \in \mathbb{R}\}. \quad (12.15)$$

- Pro $p = \infty$ minimalizujeme funkci $f(x) = \max_{i=1}^m |x - b_i|$. Řešením je střed intervalu krajních bodů, $x = \frac{1}{2}(\min_{i=1}^m b_i + \max_{i=1}^m b_i)$.
- Pro $p = 2$ minimalizujeme funkci $f(x) = \sqrt{\sum_{i=1}^m (x - b_i)^2}$. Řešením je aritmetický průměr, $x = \frac{1}{m} \sum_{i=1}^m b_i$ (viz Příklad 5.5).
- Pro $p = 1$ minimalizujeme funkci $f(x) = \sum_{i=1}^m |x - b_i|$. Řešením x je *medián* z čísel b_i (dokažte!). Medián se vypočte tak, že seřadíme čísla b_i podle velikosti a vezmeme prostřední z nich. Pokud je m sudé, máme dva 'prostřední prvky' a v tom případě funkce f nabývá minima v jejich libovolné konvexní kombinaci. Je pak úzus definovat medián jako aritmetický průměr prostředních prvků.

Dodělat důkaz, že minimum v 1-normě je medián.

Předpokládejme nyní, že jeden libovolný bod (např. b_1) se bude zvětšovat. V tom případě se řešení x pro různá p budou chovat různě. Např. aritmetický průměr se bude zvětšovat, a to tak, že zvětšováním hodnoty b_1 dosáhneme *libovolné* hodnoty x . Pro medián to ovšem neplatí – zvětšováním jediného bodu b_i ovlivníme x jen natolik, nakolik to změní pořadí bodů. Jeho libovolným zvětšováním nedosáhneme libovolné hodnoty x .

Příklad 12.12. Šuplérrou změříme průměr ocelové kuličky v několika místech, dostaneme hodnoty $\mathbf{b} = (1.02, 1.04, 0.99, 2.03)$ (cm). Při posledním měření jsme se na stupnici přehlédli, proto je poslední hodnota úplně špatně. Z těchto měření chceme odhadnout skutečný průměr. Máme

$$\frac{1}{2} \left(\min_{i=1}^m b_i + \max_{i=1}^m b_i \right) = 1.51, \quad \frac{1}{m} \sum_{i=1}^m b_i = 1.27, \quad \text{median}_{i=1}^m b_i = 1.03.$$

Je zjevné, že medián je neovlivněn vychýleným bodem, zatímco ostatní odhady ano. \square

Ve složitějším případě, např. prokládání dat polynomem jako v Příkladu 5.5, se nedá robustnost řešení ve smyslu 1-normy takto jednoduše formálně ukázat a analýza může být mnohem těžší. Ale intuitivně bude situace obdobná: řešení ve smyslu 1-normy bude méně citlivé na vychýlené body než řešení ve smyslu 2-normy.

12.6 Cvičení

12.1. Které z následujících množin jsou konvexní polyedry? Pokud je množina konvexní polyedr, dokážete ji vyjádřit ve tvaru $\{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}\}$ (tj. jako průnik poloprostorů)?

- $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0}, \sum_i x_i a_i = b, \sum_i x_i a_i^2 = c\}$, kde a_i, b, c jsou dané skaláry
- $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq \|\mathbf{x} - \mathbf{b}\|_2\}$, kde \mathbf{a}, \mathbf{b} jsou dány
- $\{\mathbf{Cx} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1\}$, kde matice \mathbf{C} je dána
- $\{\mathbf{Cx} \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_2 \leq 1\}$, kde matice \mathbf{C} je dána

12.2. Mějme vektory $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Pro každé $i = 1, \dots, m$ definujeme množinu

$$X_i = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}_i\|_2 \leq \|\mathbf{x} - \mathbf{a}_j\|_2, j \neq i\}.$$

Ukažte, že množiny X_1, \dots, X_m jsou konvexní polyedry. Ukažte, že tyto množiny tvoří rozklad (zopakujte si, co je to rozklad množiny) množiny \mathbb{R}^n . Sjednocení hranic těchto množin se nazývá *Voronoiův diagram*. Nakreslete si ho pro $n = 2$ a $m = 4$ pro různé konfigurace bodů $\mathbf{a}_1, \dots, \mathbf{a}_4$.

12.3. Najděte graficky množinu optimálních řešení úlohy

$$\begin{aligned} \min \quad & c_1 x_1 + c_2 x_2 + c_3 x_3 \\ \text{za podm.} \quad & x_1 + x_2 \geq 1 \\ & x_1 + 2x_2 \leq 3 \\ & x_1 + x_2 \leq 10 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

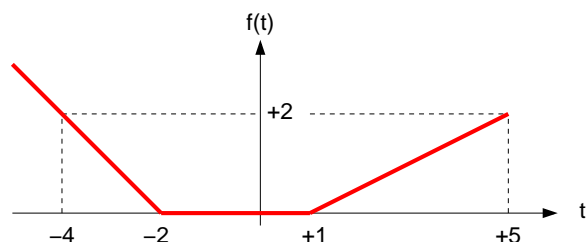
pro následující případy: $\mathbf{c} = (-1, 0, 1)$, $\mathbf{c} = (0, 1, 0)$, $\mathbf{c} = (0, 0, -1)$.

12.4. Vyřešte úvahou tyto jednoduché úlohy LP a napište co nejjednodušší vzorec pro optimální hodnotu. Vektor $\mathbf{c} \in \mathbb{R}^n$ a číslo $k \in \mathbb{N}$, $1 \leq k \leq n$, jsou dány.

- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$ (výsledek: $\sum_{i \mid c_i > 0} c_i$, tedy součet kladných čísel c_i)
- $\max\{\mathbf{c}^T \mathbf{x} \mid -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}$ (výsledek: $\sum_i |c_i|$)
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} \leq 1\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid -1 \leq \mathbf{1}^T \mathbf{x} \leq 1\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = k\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} = k\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{1}^T \mathbf{x} \leq k\}$
- $\max\{\mathbf{c}^T \mathbf{x} \mid 0 \leq x_1 \leq x_2 \leq \dots \leq x_n \leq 1\}$ (nápopověda: substituujte $y_i = x_i - x_{i-1}$)

12.5. Převeďte na LP nebo odůvodněte, proč to nejde.

- a) $\max\{|x_1 - 1| + 2|x_2 + 1| \mid x_1 + x_2 \leq 2\}$
 b) $\min\{|x_1| + |x_2| + |x_3| \mid 2x_1 - x_2 - x_3 \geq 1, -x_1 + 2x_2 - x_3 \geq 1, -x_1 - x_2 + 2x_3 \geq 1\}$
 c) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$
 d) $\max\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, |\mathbf{d}^T \mathbf{x} + e| \leq f, \mathbf{x} \geq \mathbf{0}\}$
 e) $\min\left\{\sum_{l=1}^L \max_{k=1}^K (\mathbf{c}_{kl}^T \mathbf{x} + d_{kl}) \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\right\}$
 f) $\min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^m f(\mathbf{a}_i^T \mathbf{x} - b_i)$, kde funkce f je definována obrázkem



- g) $\min\{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\|_\infty \leq 1\}$
 h) $\min\{\|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{b}\}$
 i) $\min\{\|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty \leq 1\}$
 j) $\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 + \|\mathbf{x}\|_\infty)$

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) $\max\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| = 1\} = \max\{\|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n, \|\mathbf{x}\| \leq 1\}$

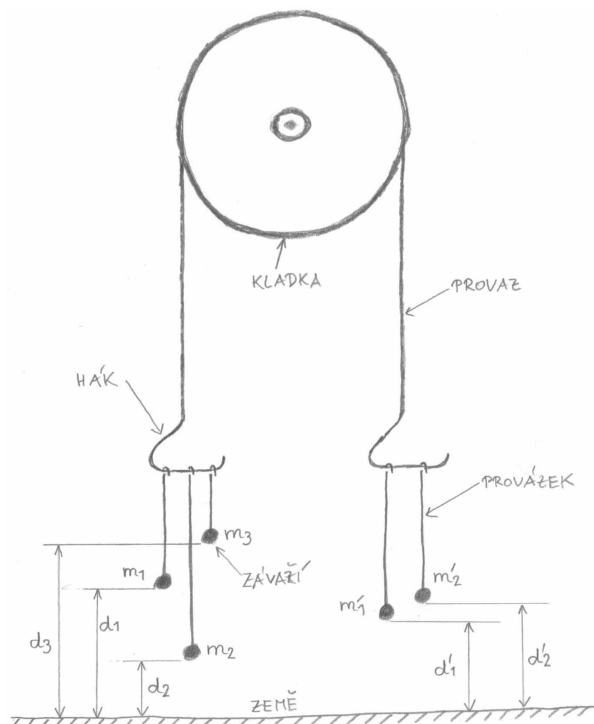
Nápověda: Inspirujte se úvahou v §12.3.1.

12.7. Pochopte kód v Matlabu, který nakreslí graf funkce $f(\mathbf{x}) = \max_{k=1}^K (\mathbf{c}_k^T \mathbf{x} + d_k)$ pro $n = 2$:

```
K = 200; N = 40;
cd = randn(3,K);
x1 = ones(N,1)*linspace(-1,1,N); x2 = linspace(-1,1,N)'*ones(1,N);
x = [x1(:)'; x2(:)']; x(3,:) = 1;
meshc(x1,x2,reshape(max(cd'*x, [],1), [N N])); axis vis3d
```

12.8. Hledáme největší hyperkouli $B(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq r\}$, která se vejde do polyedru $P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$. Tedy hledáme maximální r za podmínky $B(\mathbf{a}, r) \subseteq P$, kde optimalizujeme přes proměnné (\mathbf{a}, r) . Vyjádřete jako LP.

12.9. Máme kladku s provazem, jehož oba konce končí hákem. Na levém háku visí n závaží na provázcích, přičemž i -té závaží má tíhu m_i a jeho výška nad zemí je d_i , pro $i = 1, \dots, n$. Na pravém háku visí n' závaží na provázcích, přičemž i -té závaží má tíhu m'_i a jeho výška nad zemí je d'_i , pro $i = 1, \dots, n'$. Výšky d_i a d'_i se měří v poloze, kdy jsou oba háky ve stejné výšce nad zemí. Kladka se pohybuje bez tření, provaz a provázky jsou nekonečně ohebné, provázky a háky mají nulovou hmotnost. Obrázek ukazuje příklad pro $n = 3$, $n' = 2$.



Soustava má jediný stupeň volnosti daný otáčením kladky. Označme jako x výšku levého háku nad bodem, kdy jsou oba háky ve stejné výšce – tedy pro $x = 0$ jsou oba háky ve stejné výšce a pro $x > 0$ bude levý hák o $2x$ výše než pravý hák. V závislosti na x každé závaží buď visí nad zemí (pak je jeho potenciální energie rovna m_i krát výška nad zemí) nebo leží na zemi (pak je jeho potenciální energie nulová). Soustava bude v rovnováze při minimální celkové potenciální energii.

Napište vzorec pro celkovou potenciální energii soustavy jako funkci x . Je-li to možné, napište lineární program, jehož optimum je rovno minimální potenciální energii soustavy. Není-li to možné, vysvětlete.

12.10. Veverka před zimou potřebuje přerovnat zásoby oříšků. Stávající zásoby má v m jamkách, přičemž i -tá jamka má souřadnice $\mathbf{p}_i \in \mathbb{R}^2$ a je v ní a_i oříšků. Potřebuje je přenosit do n nových připravených jamek, přičemž j -tá jamka má souřadnice $\mathbf{q}_j \in \mathbb{R}^2$ a na konci v ní bude y_j oříšků. Veverka unese najednou jen jeden oříšek. Necht' x_{ij} označuje celkový počet oříšků přenesených ze staré jamky i do nové jamky j . Uvažujte dvě úlohy:

- Čísla y_j jsou dána. Hledají se taková čísla x_{ij} , aby se veverka vykonala co nejméně práce, kde práce na přenesení jednoho oříšku je přímo úměrná vzdálenosti (vzdušnou čarou). Běh bez oříšku se za práci nepovažuje.
- Hledají se čísla x_{ij} a y_j tak, aby veverka vykonala co nejméně práce a navíc byly v nových jamkách oříšky rozloženy co nejrovnoměrněji, čímž minimalizuje škodu způsobenou případnou krádeží. Přesněji, aby rozdíl mezi největším a nejmenším z čísel y_j byl menší než dané číslo t .

Formulujte obě úlohy jako LP. Předpokládejte, že počty oříšků jsou nezáporná reálná čísla, ač ve skutečnosti mohou být pouze nezáporná celá čísla.

Kapitola 13

Simplexová metoda

Motivace: Pokud lineární funkce na konvexním polyedru $\{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}\}$ nabývá minima, minimum se nabývá alespoň v jednom vrcholu. Naivní algoritmus na řešení LP: enumerujte všechny vrcholy tak, že zkoušíte všechny podmnožiny aktivních omezení. Aktivní omezení vyřešíte jako soustavu rovnic, přičemž chceme aby tato soustava ostatní měla právě jedno řešení. Toto řešení pak jen zkontrolujete na ostatní omezení. Ukázat příklad: 4 nerovnice v \mathbb{R}^2 takový, že nějaká množina aktivních omezení má jediné řešení, které ale neleží v polyedru.

Problém: polyedr může mít počet vrcholů exponenciální v počtu nerovnic; příkladem je hyperkrychle. Simplexová metoda přechází od vrcholu k vrcholu, přičemž nezhoršuje účelovou funkci.

George Dantzig 1947. Jeden z 10 nejslavnějších numerických algoritmů 20. století (SIAM top 10 algorithms).

Pozor: Simplexová metoda se hodně těžko přednáší. Je nutno ji nevysvětlovat obecně, ale pouze na příkladech. Obecný popis až ve skriptech. Je nutno napsat jeddu tabulku a tu postupně měnit a všechno na ní vysvětlit. Psát indexy i, i', j nalevo a nad tabulku. Je dobře si udělat předem přípravu na papír.

Zde popíšeme algoritmus na řešení úloh lineárního programování zvaný **simplexová metoda**.

Zapomeňme prozatím na účelovou funkci a zkoumejme množinu přípustných řešení LP ve tvaru

$$X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}, \quad (13.1)$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ je široká ($m < n$) matice s hodnotí m , tedy její řádky jsou lineárně nezávislé.

Soustava $\mathbf{Ax} = \mathbf{b}$ má nekonečně mnoho řešení. Položíme-li však $n - m$ složek vektoru \mathbf{x} rovno nule (tedy učiníme-li $n - m$ z podmínek $\mathbf{x} \geq \mathbf{0}$ aktivních), soustava má nejvýše jedno řešení. Tato úvaha vede k následujícím definicím:

- Množina $J \subseteq \{1, 2, \dots, n\}$ se nazývá **báze** úlohy, pokud $|J| = m$ a sloupce matice \mathbf{A} s indexy J jsou lineárně nezávislé. Tedy sloupce J tvoří regulární matici $m \times m$.
- Vektor \mathbf{x} je **bázové řešení** příslušné bázi J , pokud $\mathbf{Ax} = \mathbf{b}$ a $x_j = 0$ pro $j \notin J$.
- Bázové řešení \mathbf{x} je **přípustné**, pokud $\mathbf{x} \geq \mathbf{0}$.
- Bázové řešení \mathbf{x} je **degenerované**, pokud má méně než m nenulových složek.
- Dvě báze jsou **sousední**, pokud mají $m - 1$ společných prvků.

Protože matice \mathbf{A} má hodnot m , existuje aspoň jedna báze a každé bázi přísluší právě jedno bázové řešení. Bázové řešení však může příslušet více než jedné bázi, což se stane právě tehdy,

když je toto bázové řešení degenerované.

Příklad 13.1. Necht' je soustava $\mathbf{Ax} = \mathbf{b}$ dána tabulkou (blokovou maticí)

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} -1 & 1 & 3 & 1 & 0 & 2 & 1 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \\ -1 & 0 & 4 & 1 & 1 & 4 & 2 \end{array} \right]. \quad (13.2)$$

- $J = \{2, 3, 5\}$ není báze, protože sloupce 2, 3, 5 matice \mathbf{A} jsou lineárně závislé.
- $J = \{1, 4, 5\}$ je báze, protože tyto sloupce jsou lineárně nezávislé. Bázové řešení $\mathbf{x} = (x_1, x_2, \dots, x_6)$ příslušné bázi J se najde řešením soustavy

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \quad (13.3)$$

a položením $x_2 = x_3 = x_6 = 0$. Soustava (13.3) má právě jedno řešení, neboť její matice je regulární. Dostaneme $\mathbf{x} = (3, 0, 0, 4, 1, 0)$. Toto bázové řešení je přípustné. Není degenerované, protože má $m = 3$ nenulových složek.

- $J = \{1, 2, 4\}$ je báze. Bázové řešení je $\mathbf{x} = (4, -1, 0, 6, 0, 0)$. Je nepřípustné, protože $x_2 < 0$.
- $J = \{3, 4, 5\}$ je báze. Bázové řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$ je degenerované, protože má méně než $m = 3$ nenulových složek.
- Stejně bázové řešení $\mathbf{x} = (0, 0, 1, -2, 0, 0)$ dostaneme volbou báze $J = \{3, 4, 6\}$. Vidíme, že degenerované bázové řešení odpovídá více než jedné bázi.
- Báze $\{1, 4, 5\}$ a $\{2, 4, 5\}$ jsou sousední, protože mají společné dva prvky $\{4, 5\}$. Báze $\{1, 4, 5\}$ a $\{2, 4, 6\}$ nejsou sousední. \square

Následující věta udává spojitost mezi algebraickým a geometrickým popisem konvexního polyedru (13.1). Důkaz vynecháme, není obtížný.

Věta 13.1. *Přípustná bázová řešení jsou vrcholy polyedru (13.1), přičemž dvojice sousedních bází odpovídají dvojici vrcholů spojených hranou.*

Zde je problém: co když se optimální řešení LP nabývá na stěně (tedy úloha bude mít konečné řešení), ale tato stěna nemá žádný vrchol? Nutno vysvětlit lépe.

Věta 12.2 říká, že všechna minima lineární funkce na konvexním polyedru X se nabývají na nějaké stěně polyedru. Předpokládejme, že tato optimální stěna obsahuje alespoň jeden vrchol. Pak se alespoň jedno minimum nabývá i v tomto vrcholu. To nám dovoluje navrhnout algoritmus na řešení LP 'hrubou silou': uděláme výčet všech přípustných bázových řešení a nalezneme to s nejlepší hodnotou účelové funkce. Tato metoda je jen o málo 'chytřejší' než metoda popsaná v §9.5 a nelze prakticky použít, protože bázových řešení může být příliš mnoho.

Simplexová metoda je efektivnější obměna tohoto přístupu: přechází mezi sousedními bázemi tak, že bázová řešení jsou stále přípustná (tedy přechází po hranách polyedru X) a účelová funkce se zlepšuje (nebo aspoň nezhoršuje).

13.1 Stavební kameny algoritmu

Zde vysvětlíme jednotlivé stavební kameny simplexové metody, které nakonec v §13.2 spojíme v celý algoritmus.

13.1.1 Přechod k sousední standardní bázi

Simplexová metoda pracuje pouze se *standardními* bázemi, tj. sloupce J jsou (permutované) vektory standardní báze. To má výhodu v tom, že (i) nemusíme kontrolovat, zda jsou sloupce J lineárně nezávislé a (ii) nenulové složky bázevého 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 nahradit bázevý sloupec $j' \in J$ nebázevým sloupcem $j \notin J$. Necht' i je řádek, ve kterém je $a_{ij'} = 1$. Prvek a_{ij} se nazývá **pivot** (angl. znamená *čep*). Necht' $a_{ij} \neq 0$. Chceme nastavit pivot a_{ij} na jedničku, vynulovat prvky nad i pod pivotem, a nezměnit přitom sloupce $J \setminus \{j'\}$. Toho se dosáhne těmito ekvivalentními řádkovými úpravami:

1. Vyděl řádek i číslem a_{ij} .
2. Pro každé $i' \neq i$ odečti $a_{i'j}$ -násobek řádku i od řádku i' .

Říkáme, že jsme provedli *ekvivalentní úpravu kolem pivotu* s indexy (i, j) .

Příklad 13.2. Mějme soustavu

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & \boxed{1} & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array} \right] \quad (13.4)$$

se (standardní) bází $J = \{1, 4, 5\}$. Vidíme ihned odpovídající bázevé řešení, $\mathbf{x} = (3, 0, 0, 4, 1, 0)$.

Chceme nahradit bázevý sloupec $j' = 1$ nebázevým sloupcem $j = 2$, tedy přejít k sousední bázi $\{2, 4, 5\}$. Máme $i = 2$, tedy pivot je prvek a_{22} (v tabulce zvýrazněn). Ekvivalentními řádkovými úpravami musíme docílit, aby pivot byl roven jedné a prvky nad ním a pod ním byly nulové. Při tom smíme změnit sloupec 1, ale sloupce 4 a 5 se změnit nesmějí. Toho se docílí vydělením řádku 2 číslem a_{22} (což zde nemá žádný efekt, protože náhodou $a_{22} = 1$) a pak přičtením vhodných násobků řádku 2 k ostatním řádkům. Výsledek:

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|c} -2 & 0 & 0 & 1 & 0 & 0 & -2 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \end{array} \right].$$

Nyní sloupce $\{2, 4, 5\}$ tvoří standardní bázi. □

13.1.2 Kdy je sousední bázevé řešení přípustné?

Uvedeným způsobem můžeme od aktuální báze přejít k libovolné sousední bázi. Přitom nové bázevé řešení může nebo nemusí být přípustné. Je-li aktuální bázevé řešení přípustné, jak poznáme, zda i nové bázevé řešení bude přípustné?

Protože nenulové složky bázevého řešení \mathbf{x} jsou rovny složkám vektoru \mathbf{b} , bázevé řešení je přípustné právě tehdy, když $\mathbf{b} \geq \mathbf{0}$. Necht' v aktuální tabulce je $\mathbf{b} \geq \mathbf{0}$. Proved'me ekvivalentní úpravu kolem pivotu (i, j) . Hledáme podmínky na (i, j) , za kterých bude i po úpravě $\mathbf{b} \geq \mathbf{0}$.

Po ekvivalentní úpravě kolem pivotu (i, j) se vektor \mathbf{b} změní takto (viz §13.1.1):

- b_i se změní na b_i/a_{ij} ,

- pro každé $i' \neq i$ se $b_{i'}$ změní na $b_{i'} - a_{i'j}b_i/a_{ij}$.

Tato čísla musejí být nezáporná. To nastane právě tehdy, když platí následující podmínky:

$$a_{ij} > 0, \quad (13.5a)$$

$$\text{pro každé } i' \neq i \text{ platí } a_{i'j} \leq 0 \text{ nebo } \frac{b_i}{a_{ij}} \leq \frac{b_{i'}}{a_{i'j}}, \quad (13.5b)$$

kde 'nebo' je užito v nevylučovacím smyslu. Podmínka (13.5a) je zřejmá. Podmínka (13.5b) je ekvivalentní podmínce $b_{i'} - a_{i'j}b_i/a_{ij} \geq 0$, neboť $a_{ij} > 0$, $b_i \geq 0$, $b_{i'} \geq 0$ (rozmyslete!). *Tohle je dobré si zopakovat před přednáškou.*

Příklad 13.3. Uvažujme opět tabulku (13.4).

- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 2)$ nepovede k přípustnému bázevému řešení, neboť $a_{ij} = -1 < 0$, což porušuje podmínku (13.5a).
- Ekvivalentní úprava okolo pivotu $(i, j) = (2, 2)$ nepovede k přípustnému bázevému řešení, neboť pro $i' = 1$ je $a_{i'j} > 0$ a $\frac{3}{1} > \frac{4}{2}$, tedy podmínka (13.5b) je porušena.
- Ekvivalentní úprava okolo pivotu $(i, j) = (3, 6)$ povede k přípustnému bázevému řešení. Podmínky (13.5) jsou splněny, neboť $a_{ij} = 2 > 0$ a $\frac{1}{2} \leq \frac{4}{4}$, $\frac{1}{2} \leq \frac{3}{2}$. \square

13.1.3 Co když je celý sloupec nekladný?

Jestliže jsou všechny prvky v nějakém nebázevému sloupci j nekladné, tento sloupec se nemůže stát bázevým, neboť v něm nelze vybrat pivot splňující podmínku (13.5a). V tom případě se složka x_j bázevého řešení \mathbf{x} může zvětšovat nade všechny meze. Tedy existuje polopřímka s počátkem v \mathbf{x} ležící celá v polyedru X . To znamená, že polyedr X je neomezený.

Příklad 13.4. Necht' $[\mathbf{A} \ \mathbf{b}]$ je tabulka

$$\begin{array}{cccccc|c} 0 & -2 & 6 & 1 & 0 & 4 & 4 \\ 1 & -1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \\ \hline \mathbf{x} = & 3 & 0 & 0 & 4 & 1 & 0 \end{array}$$

s bází $\{1, 4, 5\}$. Pod tabulkou je napsáno bázevé řešení \mathbf{x} . Když se x_2 bude libovolně zvětšovat, změnu lze kompenzovat současným zvětšováním bázevých proměnných x_1, x_4, x_5 tak, že vektor \mathbf{Ax} zůstane nezměněn a tedy roven \mathbf{b} . Konkrétně, vektor pro každé $\alpha \geq 0$ bude vektor $\mathbf{x} = (3, 0, 0, 4, 1, 0) + \alpha(1, 1, 0, 2, 1, 0)$ splňovat $\mathbf{Ax} = \mathbf{b}$ a $\mathbf{x} \geq \mathbf{0}$. \square

13.1.4 Ekvivalentní úpravy účelového řádku

Dosud jsme prováděli ekvivalentní řádkové úpravy pouze na soustavě $\mathbf{Ax} = \mathbf{b}$ a účelové funkce si nevšímalí. Tyto úpravy lze rozšířit na celou úlohu LP včetně účelové funkce. Nebudeme účelovou funkci uvažovat ve tvaru $\mathbf{c}^T \mathbf{x}$, ale v mírně obecnějším tvaru $\mathbf{c}^T \mathbf{x} - d$. Tedy řešíme LP

$$\min\{\mathbf{c}^T \mathbf{x} - d \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}. \quad (13.6)$$

Úlohu budeme reprezentovat **simplexovou tabulkou**

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix}. \quad (13.7)$$

Přičtíme k účelovému řádku $[\mathbf{c}^T \ d]$ libovolnou lineární kombinaci $\mathbf{y}^T[\mathbf{A} \ \mathbf{b}]$ ostatních řádků $[\mathbf{A} \ \mathbf{b}]$, kde \mathbf{y} jsou koeficienty lineární kombinace. Ukážeme, že tato úprava zachová hodnotu účelové funkce $\mathbf{c}^T \mathbf{x} - d$ pro každé \mathbf{x} splňující $\mathbf{Ax} = \mathbf{b}$. Nový účelový řádek bude

$$[\mathbf{c}^T \ d] + \mathbf{y}^T[\mathbf{A} \ \mathbf{b}] = [\mathbf{c}^T + \mathbf{y}^T \mathbf{A} \ d + \mathbf{y}^T \mathbf{b}].$$

Nová účelová funkce bude tedy

$$(\mathbf{c}^T + \mathbf{y}^T \mathbf{A})\mathbf{x} - (d + \mathbf{y}^T \mathbf{b}) = \mathbf{c}^T \mathbf{x} - d + \mathbf{y}^T(\mathbf{Ax} - \mathbf{b}).$$

Ale to je rovno $\mathbf{c}^T \mathbf{x} - d$ pro každé \mathbf{x} splňující $\mathbf{Ax} = \mathbf{b}$.

13.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?

Necht' J je standardní báze. 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 13.5. Mějme úlohu se simplexovou tabulkou

$$\begin{bmatrix} \mathbf{c}^T & d \\ \mathbf{A} & \mathbf{b} \end{bmatrix} = \left[\begin{array}{cccccc|c} 1 & -2 & -3 & -1 & 2 & 1 & 4 \\ 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array} \right],$$

kde $J = \{1, 4, 5\}$. Složky vektoru \mathbf{c} v 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 $c_2 = 1 > 0$. \square

Pokud v některém sloupci j je $c_j \leq 0$ a $a_{ij} \leq 0$ pro všechna i , pak můžeme proměnnou x_j libovolně zvětšovat (viz §13.1.3) a účelovou funkci libovolně zmenšovat. Úloha je tedy neomezená.

13.2 Základní algoritmus

Spojením popsaných stavebních kamenů dostaneme iteraci simplexového algoritmu na řešení úlohy (13.6). Iterace přejde k sousední standardní bázi takové, že bázové řešení zůstane přípustné a účelová funkce se zmenší nebo alespoň nezmění. Vstupem i výstupem iterace je simplexová tabulka (13.7) s těmito vlastnostmi:

- podmnožina sloupců \mathbf{A} tvoří standardní bázi J ,
- bázové řešení odpovídající této bázi je přípustné, $\mathbf{b} \geq \mathbf{0}$,
- složky vektoru \mathbf{c} v bázových sloupcích jsou nulové, $c_j = 0$ pro $j \in J$.

Iteraci se provede v těchto krocích:

1. Vyber index j pivotu tak, aby $c_j < 0$ (viz §13.1.5).
2. Vyber index i pivotu podle podmínek (13.5). Z těchto podmínek plyne (promyslete!)

$$i \in \operatorname{argmin}_{i' | a_{i'j} > 0} \frac{b_{i'}}{a_{i'j}}, \quad (13.8)$$

kde $\operatorname{argmin}_{i' | a_{i'j} > 0}$ označuje, že se minimalizuje přes všechna i' splňující $a_{i'j} > 0$.

3. Udělej ekvivalentní úpravu okolo pivotu (i, j) (viz §13.1.1).
4. Udělej ekvivalentní úpravu účelového řádku, která vynuluje c_j v novém bázovém sloupci j (viz §13.1.5).

Algoritmus, který opakuje uvedenou iteraci, nazveme **základní simplexový algoritmus**. Algoritmus končí, když už nelze iteraci provést. To nastane z jednoho z těchto důvodů:

- Všechny koeficienty c_j jsou nezáporné. Účelovou funkci nelze zlepšit a jsme v optimu.
- V některém sloupci j je $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i . Úloha je neomezená.

Výběr indexů (i, j) pivotu v krocích 1 a 2 nemusí být jednoznačný, tedy může být více sloupců j s vhodným znaménkem c_j a více řádků i může splňovat podmínky (13.5) (tedy může být více argumentů minima v podmínce (13.8)). Algoritmus, který vybírá jediný pivot z těchto možností, se nazývá **pivotové pravidlo**.

Zřídka se algoritmus může dostat do stavu, kdy cyklicky prochází stále stejnou množinou bází, které odpovídají jedinému degenerovanému bázovému řešení a tedy účelová funkce se nemění. Tomuto problému **cyklení** se dá zabránit použitím vhodného pivotového pravidla (nejznámější je *Blandovo anticyklické pravidlo*), které ale popisovat nebudeme.

Příklad 13.6. Vyřešte lineární program (13.6) simplexovou metodou, když výchozí simplexová tabulka (13.7) je

$$\begin{array}{cccccc|c} 0 & -2 & 1 & 0 & 0 & -3 & 0 \\ 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & \boxed{2} & 1 \end{array}.$$

Báze je $J = \{1, 4, 5\}$ a 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. Je rozumné vzít nejmenší takové číslo, zde -3 , tedy $j = 6$.

2. Vybereme řádek i pivotu dle (13.8) nalezením argumentu minima z čísel $\frac{4}{4}, \frac{3}{2}, \frac{1}{2}$. Bude tedy $i = 3$. Výsledný pivot je označen rámečkem. Všimněte si, že řádek $i = 3$ má v aktuální bázi jedničku ve sloupci 5, sloupec 5 tedy bázi opustí.
- 3, 4. Uděláme ekvivalentní úpravu okolo pivotu (i, j) a zároveň vynulujeme číslo c_j . Neboli chceme, aby se z pivotu a_{ij} stala jednička a nad i pod pivotem byly nuly, a to včetně nultého řádku. Tedy nejprve třetí řádek vydělíme dvěma a potom k nultému řádku přičteme trojnásobek třetího řádku, od prvního řádku odečteme čtyřnásobek třetího řádku, a od druhého řádku odečteme dvojnásobek třetího řádku. Všimněte si: k žádnému řádku nikdy nepřičítáme násobky jiného řádku než pivotového. Výsledek:

$$\begin{array}{cccccc|c} 0 & -3.5 & 2.5 & 0 & 1.5 & 0 & 1.5 \\ 0 & \boxed{4} & 4 & 1 & -2 & 0 & 2 \\ 1 & 2 & 2 & 0 & -1 & 0 & 2 \\ 0 & -0.5 & 0.5 & 0 & 0.5 & 1 & 0.5 \end{array}$$

Na konci první iterace máme bázi $J = \{1, 4, 6\}$, 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 (13.8) porovnáním čísel $\frac{2}{4}, \frac{2}{2}$, tedy $i = 1$. Výsledek druhé iterace:

$$\begin{array}{cccccc|c} 0 & 0 & 6 & 0.875 & -0.25 & 0 & 3.25 \\ 0 & 1 & 1 & 0.25 & -0.5 & 0 & 0.5 \\ 1 & 0 & 0 & -0.5 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0.125 & \boxed{0.25} & 1 & 0.75 \end{array}$$

Výsledek třetí iterace:

$$\begin{array}{cccccc|c} 0 & 0 & 7 & 1 & 0 & 1 & 4 \\ 0 & 1 & 3 & 0.5 & 0 & 2 & 2 \\ 1 & 0 & 0 & -0.5 & 0 & 0 & 1 \\ 0 & 0 & 4 & 0.5 & 1 & 4 & 3 \end{array}$$

Protože všechna čísla v účelovém řádku jsou nezáporná, algoritmus končí. Úloha má optimální řešení s hodnotou -4 v bodě $(x_1, x_2, x_3, x_4, x_5, x_6) = (1, 2, 0, 0, 3, 0)$. \square

Příklad 13.7. Necht' simplexová tabulka (13.7) je

$$\begin{array}{cccccc|c} -2 & 6 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 & 0 & 2 \\ \boxed{2} & -1 & -2 & 0 & 1 & 0 & 1 \end{array}$$

Tabulka po první iteraci je

$$\begin{array}{cccccc|c} 0 & 5 & -1 & 0 & 1 & 0 & 1 \\ 0 & -1.5 & -2 & 1 & 0.5 & 0 & 2.5 \\ 1 & -0.5 & -1 & 0 & 0.5 & 0 & 0.5 \end{array}$$

Podle nultého řádku by další pivot měl být ve třetím sloupci. Ale čísla a_{i3} jsou všechna záporná (viz §13.1.3). Tedy úloha je neomezená. V nové tabulce je vidět, že můžeme zvětšovat x_3 libovolně a kompenzovat to vhodným nárůstem x_1 a x_4 . Jelikož $c_1 = c_4 = 0$, změny x_1 a x_4 se na účelové funkci neprojeví a jediný vliv na ní bude mít x_3 , které ho bude libovolně zmenšovat. \square

13.3 Inicializace algoritmu

Na začátku základního simplexového algoritmu musí být úloha zadána ve tvaru

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}, \quad (13.9)$$

kde matice \mathbf{A} obsahuje standardní bázi a $\mathbf{b} \geq \mathbf{0}$. Ukážeme, jak lze obecnou úlohu LP převést na tento tvar.

Někdy je převod snadný. Pokud má úloha tvar $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ a platí $\mathbf{b} \geq \mathbf{0}$, přidáme slackové proměnné $\mathbf{u} \geq \mathbf{0}$ a omezení převedeme na $\mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b}$. Úloha tedy bude mít simplexovou tabulku

$$\begin{bmatrix} \mathbf{c}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{A} & \mathbf{I} & \mathbf{b} \end{bmatrix},$$

ve které sloupce příslušné proměnným \mathbf{u} tvoří standardní bázi.

Příklad 13.8. Vyřešte simplexovým algoritmem:

$$\begin{aligned} \min \quad & -3x_1 - x_2 - 3x_3 \\ \text{za podmíněk} \quad & 2x_1 + x_2 + x_3 \leq 2 \\ & x_1 + 2x_2 + 3x_3 \leq 5 \\ & 2x_1 + 2x_2 + x_3 \leq 6 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

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:

$$\begin{array}{ccccccc|c} -3 & -1 & -3 & 0 & 0 & 0 & 0 & 0 \\ \hline 2 & 1 & 1 & 1 & 0 & 0 & 0 & 2 \\ 1 & 2 & 3 & 0 & 1 & 0 & 0 & 5 \\ 2 & 2 & 1 & 0 & 0 & 1 & 0 & 6 \end{array}$$

□

13.3.1 Dvoufázová simplexová metoda

Pokud je úloha zadána v obecném tvaru, operacemi z §12.3 ji lze vždy převést do tvaru (13.9). Vynásobením vhodných řádků záporným číslem vždy zajistíme $\mathbf{b} \geq \mathbf{0}$, matice \mathbf{A} ale nemusí obsahovat standardní bázi. Máme dokonce vážnější problém: není vůbec jasné, zda úloha (13.9) je přípustná. V tomto případě nejdříve vyřešíme *pomocnou úlohu* LP, která najde *nějaké* (ne nutně optimální) přípustné řešení. Z něj pak získáme standardní bázi. Pomocná úloha je

$$\min\{\mathbf{1}^T \mathbf{u} \mid \mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{0}\} \quad (13.10)$$

a má simplexovou tabulku

$$\begin{bmatrix} \mathbf{0} & \mathbf{1}^T & 0 \\ \mathbf{A} & \mathbf{I} & \mathbf{b} \end{bmatrix}.$$

Pro libovolné $\mathbf{u} \geq \mathbf{0}$ je $\mathbf{1}^T \mathbf{u} \geq 0$, přičemž $\mathbf{1}^T \mathbf{u} = 0$ právě tehdy, když $\mathbf{u} = \mathbf{0}$. Tedy úloha (13.9) je přípustná právě tehdy, je-li optimální hodnota úlohy (13.10) rovna 0. Na počátku tvoří sloupce příslušné proměnným \mathbf{u} standardní bázi, lze tedy na ní pustit základní simplexový algoritmus. Ten může skončit dvěma způsoby:

- Pokud je optimum větší než 0, pak úloha (13.9) je nepřípustná.
- Pokud je optimum rovno 0, pak úloha (13.9) je přípustná. Pokud není optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (13.10) degenerované, po skončení simplexového algoritmu jsou všechny bázevé proměnné kladné. Protože $\mathbf{u} = \mathbf{0}$, proměnné \mathbf{u} budou tedy nebázevé. Proto mezi sloupci příslušnými proměnným \mathbf{x} bude existovat standardní báze.

Pokud je optimální řešení (\mathbf{x}, \mathbf{u}) úlohy (13.10) degenerované, některé proměnné \mathbf{u} mohou být na konci algoritmu bázevé. Pak je nutno udělat dodatečné úpravy kolem pivotů ve sloupcích příslušných bázevým proměnným \mathbf{u} , abychom tyto sloupce dostali z báze ven.

Nalezení nějakého přípustného řešení v pomocné úloze (13.10) se nazývá **první fáze** a řešení původní úlohy pak **druhá fáze** algoritmu, mluvíme tedy o **dvoufázové simplexové metodě**.

Příklad 13.9. Řešte

$$\begin{aligned} \min \quad & -20x_1 - 30x_2 - 40x_3 \\ \text{za podmíněk} \quad & 3x_1 + 2x_2 + x_3 = 10 \\ & x_1 + 2x_2 + 2x_3 = 15 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Máme sice $\mathbf{b} \geq \mathbf{0}$, ale není jasné, zda existuje přípustné \mathbf{x} , tím méně není vidět standardní báze. Provedeme první fázi algoritmu. Pomocná úloha bude

$$\begin{aligned} \min \quad & u_1 + u_2 \\ \text{za podmíněk} \quad & 3x_1 + 2x_2 + x_3 + u_1 = 10 \\ & x_1 + 2x_2 + 2x_3 + u_2 = 15 \\ & x_1, x_2, x_3, u_1, u_2 \geq 0 \end{aligned}$$

s tabulkou

$$\begin{array}{cccccc|c} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 1 & 0 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 0 & 15 \end{array}$$

Sloupce nad přidanými proměnnými tvoří standardní bázi, můžeme tedy na pomocnou úlohu pustit základní simplexový algoritmus. Po vynulování ceny nad bázevými proměnnými budou

kroky algoritmu vypadat takto:

$$\begin{array}{ccccc|c} -4 & -4 & -3 & 0 & 0 & -25 \\ \hline 3 & \boxed{2} & 1 & 1 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 15 \end{array}$$

$$\begin{array}{ccccc|c} 2 & 0 & -1 & 2 & 0 & -5 \\ \hline 1.5 & 1 & 0.5 & 0.5 & 0 & 5 \\ -2 & 0 & \boxed{1} & -1 & 1 & 5 \end{array}$$

$$\begin{array}{ccccc|c} 0 & 0 & 0 & 1 & 1 & 0 \\ \hline 2.5 & 1 & 0 & 1 & -0.5 & 2.5 \\ -2 & 0 & 1 & -1 & 1 & 5 \end{array}$$

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

$$\begin{array}{ccc|c} -20 & -30 & -40 & 0 \\ \hline 2.5 & 1 & 0 & 2.5 \\ -2 & 0 & 1 & 5 \end{array}$$

□

13.4 Cvičení

13.1. V tabulce označte všechny takové pivoty, že ekvivalentní úprava kolem nich povede k přípustnému bázovému řešení:

$$[\mathbf{A} \ \mathbf{b}] = \left[\begin{array}{cccccc|ccc} 0 & 2 & 6 & 1 & 0 & -4 & 3 & 0 & 4 \\ 1 & 1 & -3 & 0 & 0 & 2 & 3 & 0 & 3 \\ 0 & -1 & 1 & 0 & 1 & -2 & -3 & 0 & 1 \\ 0 & -2 & 2 & 0 & 0 & 2 & -1 & 1 & 1 \end{array} \right]$$

13.2. Zapište lineární program

$$\begin{array}{ll} \min & -x_1 \qquad \qquad \qquad - x_4 - 3x_5 \\ \text{za podmíněk} & 2x_1 \qquad \qquad \qquad + x_4 + x_5 + x_6 = 2 \\ & -x_1 + x_2 \qquad \qquad + 2x_4 + 3x_5 = 4 \\ & 2x_1 \qquad \qquad + x_3 + 2x_4 - x_5 = 6 \\ & \qquad \qquad \qquad \qquad \qquad \qquad x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{array}$$

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.

13.3. Vyřešte simplexovou metodou:

$$\begin{array}{ll} \max & 2x_1 - x_2 - 3x_3 \\ \text{za podmíněk} & -2x_1 - x_2 + x_3 \leq 2 \\ & -x_1 + 2x_2 - 3x_3 \leq 5 \\ & -2x_1 - 4x_2 + x_3 \leq 6 \\ & \qquad \qquad \qquad x_1, x_2, x_3 \geq 0 \end{array}$$

13.4. Vyřešte simplexovou metodou (navzdory tomu, že lze řešit úvahou):

$$\begin{aligned} \max \quad & 6x_1 + 9x_2 + 5x_3 + 9x_4 \\ \text{za podmíněk} \quad & x_1 + x_2 + x_3 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

13.5. Necht' úloha (13.6) má více než jedno optimální řešení. Jak se to pozná v simplexové tabulce? Navrhněte algoritmus, jehož výstupem bude výčet všech optimálních bázových řešení.

13.6. Upravte do vhodného tvaru a vyřešte dvoufázovou simplexovou metodou:

$$\begin{aligned} \max \quad & 3x_1 - 4x_2 \\ \text{za podmíněk} \quad & -2x_1 - 5x_2 \leq 10 \\ & 3x_1 + x_2 \leq 3 \\ & -2x_1 + x_2 \leq -2 \\ & x_1 \geq 0 \\ & x_2 \leq -1 \end{aligned}$$

Řešení: Optimum je $(x_1, x_2) = (25, -36)/13$.

Kapitola 14

Dualita v lineárním programování

Ke každé úloze LP lze sestavit podle dále popsaného postupu jinou úlohu LP. Novou úlohu nazýváme **duální**, původní úlohu nazýváme **primární** či **přímou**. Konstrukce je symetrická: duál duálu je původní úloha. Tedy má smysl říkat, že primární a duální úloha jsou *navzájem* duální. Dvojice duálních úloh je svázána zajímavými vztahy.

14.1 Konstrukce duální úlohy

K úloze LP v obecném tvaru (viz §12.3) se duální úloha získá dle tohoto postupu:

$$\begin{array}{ll}
 \min \sum_{j \in J} c_j x_j & \max \sum_{i \in I} y_i b_i \\
 \text{za podm. } \sum_{j \in J} a_{ij} x_j = b_i & \text{za podm. } y_i \in \mathbb{R}, \quad i \in I_0 \\
 \sum_{j \in J} a_{ij} x_j \geq b_i & y_i \geq 0, \quad i \in I_+ \\
 \sum_{j \in J} a_{ij} x_j \leq b_i & y_i \leq 0, \quad i \in I_- \\
 x_j \in \mathbb{R} & \sum_{i \in I} y_i a_{ij} = c_j, \quad j \in J_0 \\
 x_j \geq 0 & \sum_{i \in I} y_i a_{ij} \leq c_j, \quad j \in J_+ \\
 x_j \leq 0 & \sum_{i \in I} y_i a_{ij} \geq c_j, \quad j \in J_-
 \end{array}$$

V levém sloupci je primární úloha, v prostředním sloupci je z ní vytvořená duální úloha. V pravém sloupci jsou množiny indexů pro obě úlohy: $I = \{1, \dots, m\} = I_0 \cup I_+ \cup I_-$ je indexová množina primárních omezení a duálních proměnných, $J = \{1, \dots, n\} = J_0 \cup J_+ \cup J_-$ je indexová množina primárních proměnných a duálních omezení.

Všimněte si následující symetrie: i -tému primárnímu omezení $\sum_j a_{ij} x_j \geq b_i$ odpovídá duální proměnná $y_i \geq 0$. Opačně, j -tá primární proměnná $x_j \geq 0$ odpovídá j -tému duálnímu omezení $\sum_i y_i a_{ij} \leq c_j$. Podobně pro ostatní řádky.

Pro speciální tvary LP se dvojice duálních úloh přehledněji napíše v maticové formě. Např. pro $I_0 = I_- = J_0 = J_- = \emptyset$ obdržíme

$$\begin{array}{ll}
 \min \mathbf{c}^T \mathbf{x} & \max \mathbf{y}^T \mathbf{b} \\
 \text{za podm. } \mathbf{Ax} \geq \mathbf{b} & \text{za podm. } \mathbf{y} \geq \mathbf{0} \\
 \mathbf{x} \geq \mathbf{0} & \mathbf{y}^T \mathbf{A} \leq \mathbf{c}^T
 \end{array} \tag{14.1}$$

14.2 Věty o dualitě

Následující věty platí pro obecný tvar LP, ale důkazy uděláme pouze pro speciální tvar (14.1).

Věta 14.1 (o slabé dualitě). *Nechť \mathbf{x} je přípustné primární řešení a \mathbf{y} přípustné duální řešení. Pak $\mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T \mathbf{b}$.*

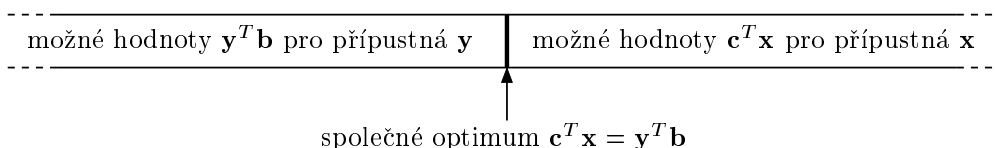
Důkaz. Díky přípustnosti \mathbf{x} a \mathbf{y} platí $\mathbf{y}^T \mathbf{A} \leq \mathbf{c}^T$ a $\mathbf{x} \geq \mathbf{0}$, z čehož plyne (proč?) $\mathbf{y}^T \mathbf{A} \mathbf{x} \leq \mathbf{c}^T \mathbf{x}$. Podobně, díky přípustnosti \mathbf{x} a \mathbf{y} platí $\mathbf{A} \mathbf{x} \geq \mathbf{b}$ a $\mathbf{y} \geq \mathbf{0}$, z čehož plyne $\mathbf{y}^T \mathbf{A} \mathbf{x} \geq \mathbf{y}^T \mathbf{b}$. Z toho

$$\mathbf{c}^T \mathbf{x} \geq \mathbf{y}^T \mathbf{A} \mathbf{x} \geq \mathbf{y}^T \mathbf{b}. \quad (14.2)$$

□

Věta 14.2 (o silné dualitě). *Primární úloha má optimální řešení právě tehdy, když má duální úloha optimální řešení. Má-li primární úloha optimální řešení \mathbf{x} a duální úloha optimální řešení \mathbf{y} , platí $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b}$.*

Důkaz věty o silné dualitě není jednoduchý a vynecháme jej. Věty o slabé a silné dualitě mají jasnou interpretaci: pro přípustná \mathbf{x} a \mathbf{y} není hodnota duální účelové funkce nikdy větší než hodnota primární účelové funkce a tyto hodnoty se potkají ve společném optimu:



Věta 14.3 (o komplementaritě). *Nechť \mathbf{x} je přípustné primární řešení a \mathbf{y} přípustné duální řešení. Pak $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b}$ právě tehdy, když zároveň platí tyto dvě podmínky:*

$$\text{Pro každé } i \in I \text{ platí } y_i = 0 \text{ nebo } \sum_{j \in J} a_{ij} x_j = b_i. \quad (14.3a)$$

$$\text{Pro každé } j \in J \text{ platí } x_j = 0 \text{ nebo } \sum_{i \in I} y_i a_{ij} = c_j. \quad (14.3b)$$

‘Nebo’ je zde užito v nevylučovacím smyslu, tj. mohou nastat obě možnosti současně.

Důkaz. Pro libovolné vektory $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ platí (rozmyslete!)

$$\forall i (u_i = 0 \text{ nebo } v_i = 0) \iff \forall i (u_i v_i = 0) \iff \mathbf{u}^T \mathbf{v} = 0.$$

Tedy podmínky (14.3) je možno psát jako

$$\mathbf{y}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) = 0, \quad (14.4a)$$

$$(\mathbf{c}^T - \mathbf{y}^T \mathbf{A}) \mathbf{x} = 0. \quad (14.4b)$$

Tvrzení $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{b}$ je ekvivalentní tomu, že obě nerovnosti v (14.2) jsou aktivní (platí s rovnostmi). Pak $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{x}$ je ekvivalentní (14.4b) a $\mathbf{y}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{b}$ je ekvivalentní (14.4a). □

Příklad 14.1. Mějme dvojici navzájem duálních úloh LP:

$$\begin{array}{ll}
 \min & 2x_1 + 5x_2 + 6x_3 = \mathbf{5.4} \\
 \mathbf{3} = & 2x_1 + x_2 + 2x_3 \geq 3 \\
 \mathbf{2.4} = & x_1 + 2x_2 + 2x_3 \geq 1 \\
 \mathbf{3} = & x_1 + 3x_2 + x_3 \geq 3 \\
 -\mathbf{0.6} = & -x_1 + x_2 - 2x_3 \geq -1 \\
 \mathbf{1.2} = & x_1 \geq 0 \\
 \mathbf{0.6} = & x_2 \geq 0 \\
 \mathbf{0} = & x_3 \geq 0 \\
 \max & 3y_1 + y_2 + 3y_3 - y_4 = \mathbf{5.4} \\
 \mathbf{0.2} = & y_1 \geq 0 \\
 \mathbf{0} = & y_2 \geq 0 \\
 \mathbf{1.6} = & y_3 \geq 0 \\
 \mathbf{0} = & y_4 \geq 0 \\
 \mathbf{2} = & 2y_1 + y_2 + y_3 - y_4 \leq 2 \\
 \mathbf{5} = & y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\
 \mathbf{2} = & 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6
 \end{array}$$

Spočetli jsme optimální řešení obou úloh a dosadili tato řešení do účelových funkcí a do omezení. Hodnoty optimálních řešení $\mathbf{x}^* = (1.2, 0.6, 0)$ a $\mathbf{y}^* = (0.2, 0, 1.6)$ a hodnoty omezení a účelových funkcí v optimech jsou napsané tučně před/za rovnítky. Dle věty o silné dualitě se optima rovnají. Vezmeme-li libovolný řádek (kromě účelového), je na něm alespoň jedno z obou omezení aktivní. Např. ve druhém řádku je primární omezení $2x_1 + x_2 + 2x_3 \geq 3$ aktivní a duální omezení $y_1 \geq 0$ je neaktivní. Podle věty o komplementaritě se nemůže stát, že by na některém řádku byly obě omezení zároveň neaktivní (mohou být obě ale zároveň aktivní, což zde nenastává, ale může to nastat v případě degenerace). \square

Zopakujme (viz §12), že pro každou úlohu LP mohou nastat 3 možnosti: úloha má optimální řešení, úloha je neomezená, úloha je nepřipustná.

Věta 14.4. 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 $-\infty \geq +\infty$. \square

Předložíme-li přípustná primární a duální řešení taková, že se účelové funkce rovnají, dokázali jsme optimalitu obou úloh. Pro velké úlohy to může být nejsnadnější důkaz optimality.

Máme-li duální optimální řešení, jak z něj co nejlevněji spočítat primární optimální řešení? Obecně je k tomu nutno vyřešit soustavu lineárních nerovnic (což není snadnější než vyřešit lineární program). Někdy ale postačí vyřešit soustavu rovnic.

Příklad 14.2. Je dána primární úloha z Příkladu 14.1. Dokažte bez použití algoritmu na řešení LP, že $\mathbf{x} = (x_1, x_2, x_3) = (1.2, 0.6, 0)$ je optimální řešení primární úlohy (přičemž není zadáno duální řešení \mathbf{y})

Optimalitu daného \mathbf{x} zkusíme dokázat pomocí věty o komplementaritě. Předpokládejme, že \mathbf{y} (které zatím neznáme) je optimální řešení duální úlohy. Protože jsou druhé a čtvrté primární

omezení neaktivní (neplatí v nich rovnost ale pouze nerovnost), z komplementarity musí být $y_2 = y_4 = 0$. Protože $x_1 > 0$ a $x_2 > 0$, z komplementarity musí být první a druhé duální omezení aktivní. Máme tedy soustavu lineárních rovnic

$$\begin{aligned} 2y_1 + y_3 &= 2 \\ y_1 + 3y_3 &= 5 \end{aligned} \tag{14.5}$$

která 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. Lze ukázat, že pokud by bylo primární optimální řešení degenerované, duální úloha by měla nekonečně mnoho optimálních. V tom případě by soustava (14.5) měla nekonečně mnoho řešení. Mezi nimi by sice byly přípustná duální řešení, ale k jejich nalezení bychom museli řešit soustavu rovnic a nerovnic. \square

14.3 Stínové ceny

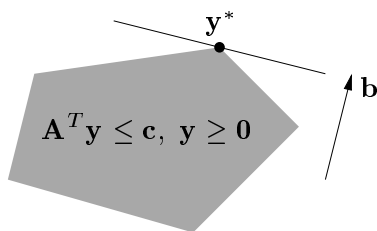
Zkoumejme, jak se změní optimální hodnota úlohy $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, jestliže nepatrně změníme pravé strany omezení \mathbf{b} . Odpověď je snadno vidět v duálu.

Věta 14.5 (o stínových cenách). *Necht' funkce $f: \mathbb{R}^m \rightarrow \mathbb{R}$ je definována jako*

$$f(\mathbf{b}) = \min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} = \max\{\mathbf{y}^T \mathbf{b} \mid \mathbf{y}^T \mathbf{A} \leq \mathbf{c}^T, \mathbf{y} \geq \mathbf{0}\}.$$

Jestliže má duální úloha pro dané \mathbf{b} jediné optimální řešení \mathbf{y}^ , pak je funkce f v bodě \mathbf{b} diferencovatelná a platí $f'(\mathbf{b}) = \mathbf{y}^{*T}$, neboli $\partial f(\mathbf{b})/\partial b_i = y_i^*$.*

Důkaz. Jelikož je duální optimální řešení \mathbf{y}^* jediné, nabývá se ve vrcholu polyedru přípustných řešení $\{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y}^T \mathbf{A} \leq \mathbf{c}^T, \mathbf{y} \geq \mathbf{0}\}$, 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 $\mathbf{y}^{*T} \mathbf{b}$. Derivací získáme $f'(\mathbf{b}) = \mathbf{y}^{*T}$. \square

Uvědomte si nutnost předpokladu o jednoznačnosti optimálního řešení. Kdyby množina duálních optimálních řešení byla ne jediný vrchol, ale stěna vyšší dimenze, po infinitezimální změně účelového vektoru \mathbf{b} by se optimální stěna mohla stát vrcholem a funkce f by tedy v bodě \mathbf{b} nebyla diferencovatelná. Předpoklad o jednoznačnosti řešení lze vypustit, ale pak by věta byla složitější.

Protože \mathbf{b} je zároveň vektor pravých stran primární úlohy, optimální duální proměnné \mathbf{y}^* vyjadřují citlivost optima primární úlohy na změnu pravých stran primárních omezení $\mathbf{Ax} \geq \mathbf{b}$. Interpretujeme-li naše LP jako optimální výrobní plán (12.6) (pozor, liší se obrácenou nerovností

v omezení), pak hodnota y_i^* říká, jak by se náš výtěžek zvětšil, kdybychom trochu uvolnili omezení na výrobní zdroje $\mathbf{a}_i^T \mathbf{x} \leq b_i$. V ekonomii se proto duálním proměnným říká **stínové ceny** primárních omezení.

Všimněte si, že věta o stínových cenách je ve shodě s větou o komplementaritě. Pokud $y_i^* = 0$, je $\mathbf{a}_i^T \mathbf{x} < b_i$, tedy malá změna b_i nemá na optimum vliv.

Příklad 14.3. Lze ukázat, že v Příkladu 14.1 má duální úloha 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* $\mathbf{y}^{*T} \mathbf{b}$. Podle silné duality hodnota primárního optima musí být rovna hodnotě duálního optima (argument \mathbf{x}^* primárního optima se nějak změní, my ale nepotřebujeme vědět, jak). Dvojice úloh tedy bude vypadat takto:

$$\begin{array}{ll}
 \min & 2x_1 + 5x_2 + 6x_3 = \mathbf{5.402} \\
 & 2x_1 + x_2 + 2x_3 \geq 3.01 \\
 & x_1 + 2x_2 + 2x_3 \geq 1 \\
 & x_1 + 3x_2 + x_3 \geq 3 \\
 & -x_1 + x_2 - 2x_3 \geq -1 \\
 & x_1 \geq 0 \\
 & x_2 \geq 0 \\
 & x_3 \geq 0 \\
 \max & 3.01y_1 + y_2 + 3y_3 - y_4 = \mathbf{5.402} \\
 & \mathbf{0.2} = y_1 \geq 0 \\
 & \mathbf{0} = y_2 \geq 0 \\
 & \mathbf{1.6} = y_3 \geq 0 \\
 & \mathbf{0} = y_4 \geq 0 \\
 & \mathbf{2} = 2y_1 + y_2 + y_3 - y_4 \leq 2 \\
 & \mathbf{5} = y_1 + 2y_2 + 3y_3 + y_4 \leq 5 \\
 & \mathbf{2} = 2y_1 + 2y_2 + y_3 - 2y_4 \leq 6
 \end{array}$$

Věta o stínových cenách říká, že v malém okolí bodu $\mathbf{b} = (3, 1, 3, -1)$, ve kterém se nemění optimální \mathbf{y}^* , bude $f(\mathbf{b}) = \mathbf{y}^{*T} \mathbf{b}$ a tedy

$$5.402 - 5.4 = \frac{\partial f(\mathbf{b})}{\partial b_1} h = y_1 h = 0.2 \cdot 0.01. \quad \square$$

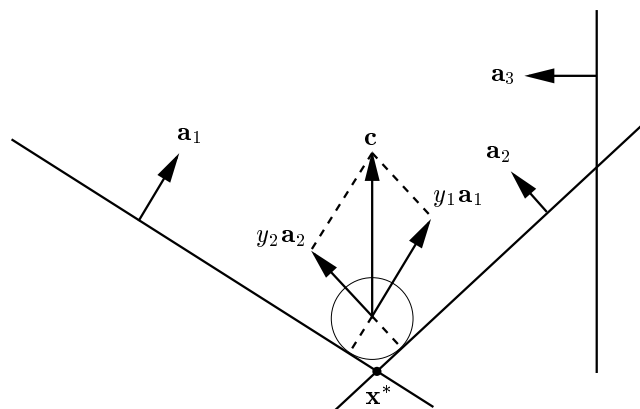
14.4 Příklady na konstrukci a interpretaci duálních úloh

Dualita umožňuje *vhled* do řešeného problému, často velmi netriviální. Abychom danou úlohu (fyzikální, ekonomickou či jinou) popsanou lineárním programem porozuměli do hloubky, je často třeba pochopit význam nejen primární úlohy, ale i duální úlohy a vět o dualitě.

Příklad 14.4 (Fyzikální interpretace duality). Uvažujme dvojici duálních úloh

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n\} = \max\{\mathbf{y}^T \mathbf{b} \mid \mathbf{y}^T \mathbf{A} = \mathbf{c}^T, \mathbf{y} \geq \mathbf{0}\}.$$

Uvažujme následující ‘analogový počítač’. Mějme polyedr tvořený třemi poloprostory $\mathbf{a}_i^T \mathbf{x} \geq b_i$ a vektor \mathbf{c} mířící svisle vzhůru:



Hodíme do polyedru malý míček, na který působí tíhová síla $-\mathbf{c}$. Míček s pozicí \mathbf{x} má potenciální energii $-\mathbf{c}^T \mathbf{x}$. Míček se bude pohybovat do té doby, než nalezne místo s nejmenší potenciální energií, což je nejnižší vrchol \mathbf{x}^* . Tedy \mathbf{x}^* je řešením primární úlohy.

V bodě \mathbf{x}^* je míček v klidu a proto pro něj platí rovnováha sil: tíha $-\mathbf{c}$ se vyrovnává silami stěn. Tedy existují skaláry $y_i^* \geq 0$ tak, že $\mathbf{c} = \sum_i y_i^* \mathbf{a}_i = \mathbf{A}^T \mathbf{y}^*$. Skaláry y_i^* jsou nezáporné, protože stěny působí silou jen dovnitř polyedru, ne ven.

Pokud $\mathbf{a}_i^T \mathbf{x}^* > b_i$, tedy míček se i -té stěny nedotýká, síla stěny je nulová, $y_i^* = 0$. Proto pro každé i platí $y_i^*(\mathbf{a}_i^T \mathbf{x}^* - b_i) = 0$ (komplementarita). Z toho máme $\mathbf{y}^{*T} \mathbf{A} \mathbf{x}^* = \mathbf{c}^T \mathbf{x}^* = \mathbf{y}^{*T} \mathbf{b}$. Tedy \mathbf{y}^* je řešením duální úlohy a platí silná dualita¹. \square

Příklad 14.5 (Ekonomická interpretace duality). Vraťme se k Příkladu 12.10 o výrobci lupínků a hranolků z brambor a oleje. Napišme k této úloze duální úlohu:

$$\begin{array}{ll} \max & 120l + 76h \\ \text{za podm.} & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l \geq 0 \\ & h \geq 0 \end{array} \qquad \begin{array}{ll} \min & 100a + 16b \\ \text{za podm.} & a \geq 0 \\ & b \geq 0 \\ & 2a + 0.4b \geq 120 \\ & 1.5a + 0.2b \geq 76 \end{array}$$

Přijde překupník a chce koupit od výrobce jeho zásoby brambor a oleje. Překupník řeší tuto otázku: Jaké nejnižší ceny musím nabídnout, aby mi výrobce své zásoby prodal? Tvrdíme, že toto je význam duální úlohy.

Vskutku, necht' a, b označují nabízenou cenu za jednotku brambor a oleje. Překupník chce minimalizovat celkovou cenu za suroviny $100a + 16b$. Musí být $2a + 0.4b \geq 120$, neboť jinak by výrobcovi více vyplatilo vyrobit ze všech brambor a oleje lupínky a prodat je, než prodat suroviny. Ze stejného důvodu musí být $1.5a + 0.2b \geq 76$. Optimální duální řešení je $a = 32$ a $b = 140$.

Toto je další důvod (kromě Věty 14.5), proč se optimálním duálním proměnným někdy říká *stínové ceny* odpovídajících primárních omezení. Např. stínová cena brambor je 32 Kč/kg. \square

Příklad 14.6. Mějme úlohu

$$\min \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0} \} = \min \left\{ \sum_{i=1}^n c_i x_i \mid \sum_{i=1}^n x_i = 1, x_i \geq 0 \right\},$$

¹ Zdůrazněme, že tato úvaha *nedokazuje* žádnou ze tří vět o dualitě. Předpokládáme totiž platnost fyzikálních zákonů, které ale nelze matematicky dokázat, lze je pouze experimentálně pozorovat.

kde $\mathbf{c} = (c_1, \dots, c_n)$ je dáno a optimalizuje se přes $\mathbf{x} = (x_1, \dots, x_n)$. Najděte elementární úvahou hodnotu optima, napište duální úlohu. Vysvětlete, co v dané úloze znamenají věty o silné dualitě a komplementaritě.

Optimální hodnota je $\min_{i=1}^n c_i$ (viz Příklad 12.4). Dosahuje se ve vektoru \mathbf{x} jehož všechny složky jsou nulové kromě složek příslušných minimálnímu c_i . To je jasné, protože je nejvýhodnější soustředit všechnu ‘váhu’ rozdělení \mathbf{x} do nejmenšího prvku. Pokud je více minimálních prvků c_i , optimální vektor \mathbf{x} není dán jednoznačně. Např. pro $\mathbf{c} = (1, 3, 1, 2)$ budou optimálními řešeními vektory $\mathbf{x} = (x_1, 0, x_3, 0)$ pro všechna $x_1, x_3 \geq 0$ splňující $x_1 + x_3 = 1$.

Podle návodu na konstrukci duální úlohy dostaneme duál

$$\max\{y \in \mathbb{R} \mid y\mathbf{1} \leq \mathbf{c}\} = \max\{y \in \mathbb{R} \mid y \leq c_i, i = 1, \dots, n\}.$$

V této úloze se hledá největší číslo y , které je menší než všechna čísla c_i . Takové číslo y se rovná minimu z čísel c_i . Tedy platí silná dualita.

Podmínky komplementarity říkají, že v optimech bude alespoň jedno z odpovídající dvojice primární-duální omezení aktivní. Dvojice omezení $\sum_i x_i = 1, y \in \mathbb{R}$ splňuje podmínky komplementarity triviálně. Dvojice omezení $x_i \geq 0, y \leq c_i$ je splňuje právě tehdy, když je splněna aspoň jedna z rovností $x_i = 0, y = c_i$. To znamená:

- Pokud je v duálu $y < c_i$, v primáru musí být $x_i = 0$. To je ale jasné, protože $y < c_i$ znamená, že c_i není nejmenší ze složek vektoru \mathbf{c} a tudíž v primáru by byla hloupost mu přiřadit nenulovou váhu x_i .
- Obráceně, pokud je v primáru $x_i > 0$, musí být v duálu $y = c_i$. To je jasné, protože pokud jsme v primáru přiřadili číslu c_i nenulovou váhu, musí být nejmenší. \square

Příklad 14.7. Víme (viz §12.5), že řešením úlohy

$$\min_{x \in \mathbb{R}} \sum_{i=1}^m |x - b_i| = \min\{\mathbf{1}^T \mathbf{z} \mid \mathbf{z} \in \mathbb{R}^m, x \in \mathbb{R}, -\mathbf{z} \leq \mathbf{1}x - \mathbf{b} \leq \mathbf{z}\}. \quad (14.6)$$

je medián čísel $\mathbf{b} = (b_1, \dots, b_m) \in \mathbb{R}^m$. Najděme k této úloze duální úlohu a výsledek co možná nejvíce zjednodušte.

Vlevo opíšeme primární úlohu a vpravo napíšeme duální úlohu:

$$\begin{array}{ll} \min & \sum_i z_i \\ \text{za podm.} & x + z_i \geq b_i \\ & -x + z_i \geq -b_i \\ & z_i \in \mathbb{R} \\ & x \in \mathbb{R} \end{array} \quad \begin{array}{ll} \max & \sum_i (p_i - q_i) b_i \\ \text{za podm.} & p_i \geq 0 \\ & q_i \geq 0 \\ & p_i + q_i = 1 \\ & \sum_i (p_i - q_i) = 0 \end{array} \quad \forall i \quad (14.7)$$

Duální úlohu jsme odvodili přesně podle návodu v §14.1. Napsat duál takto ‘přímo’ chce ovšem trochu zkušenosti.

Zdlouhavější avšak jistější cesta je přes maticovou formu. Primární a duální úlohu napíšeme ve tvaru (14.1), kde ale přejmenujeme názvy matic, aby nekolidovaly s (14.6):

$$\begin{array}{ll} \min & \mathbf{h}^T \mathbf{u} \\ \text{za podm.} & \mathbf{F} \mathbf{u} \geq \mathbf{g} \\ & \mathbf{u} \in \mathbb{R}^{1+m} \end{array} \quad \begin{array}{ll} \max & \mathbf{v}^T \mathbf{g} \\ \text{za podm.} & \mathbf{v} \geq \mathbf{0} \\ & \mathbf{v}^T \mathbf{F} = \mathbf{h}^T \end{array}$$

Matice zvolíme tak, aby primární úloha odpovídala úloze (14.6):

$$\mathbf{F} = \begin{bmatrix} 1 & \mathbf{I} \\ -1 & \mathbf{I} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} x \\ \mathbf{z} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix}.$$

Vektor duálních proměnných \mathbf{v} jsme zároveň rozdělili na dva bloky \mathbf{p}, \mathbf{q} , odpovídající blokům matic \mathbf{F} a \mathbf{g} . Vynásobním matic přepíšeme duální úlohu do tvaru (ověřte na papíře!)

$$\max\{ \mathbf{b}^T(\mathbf{p} - \mathbf{q}) \mid \mathbf{1}^T(\mathbf{p} - \mathbf{q}) = 0, \mathbf{p} + \mathbf{q} = \mathbf{1}, \mathbf{p} \geq \mathbf{0}, \mathbf{q} \geq \mathbf{0} \},$$

což ve skalárním tvaru je pravá úloha (14.7).

Duální úlohu lze dále zjednodušit substitucí

$$2p_i = 1 + t_i, \quad 2q_i = 1 - t_i.$$

Po této substituci je $p_i - q_i = t_i$, 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} má tvar

$$\max \left\{ \sum_{i=1}^m b_i t_i \mid \sum_{i=1}^m t_i = 0, -1 \leq t_i \leq 1 \right\} = \max\{ \mathbf{b}^T \mathbf{t} \mid \mathbf{1}^T \mathbf{t} = 0, -\mathbf{1} \leq \mathbf{t} \leq \mathbf{1} \}. \quad (14.8)$$

Zde vůbec není na první pohled vidět, že optimální hodnoty úloh (14.6) a (14.8) jsou stejné (silná dualita). Jako cvičení to chytrý student může zkusit odůvodnit. \square

Řešení: Podmínka $\sum_i t_i = 0$ znamená, že $\sum_i (b_i - \alpha)t_i = \sum_i b_i t_i$, tedy optimální hodnota je invariantní vůči translaci bodů. Zvolme $\alpha = \text{median}_i b_i$. Pak je zjevné, že optimální řešení je $t_i = -1$ pro $b_i < \alpha$ a $t_i = 1$ pro $b_i > \alpha$.

Napsal jsem duál k transportní úloze a vysvětlil na zařízení s vidlicemi. Nemělo to ale velký úspěch.

14.5 Cvičení

14.1. Ukažte pro dvojici úloh LP v §14.1, že duál duálu se rovná původní úloze. Musíte nejdříve duální úlohu (prostřední sloupec) vpravo převést do tvaru primární úlohy (první sloupec), tj. např. musíte převést maximalizaci na minimalizaci.

14.2. Napište duální úlohy a podmínky komplementarity k následujícím úlohám. Výsledek co nejvíce zjednodušte příp. převed'te do skalární formy, je-li skalární forma výstižnější.

- $\min_{x \in \mathbb{R}} \max_{i=1}^n |a_i - x|$ (střed intervalu)
- úloha (12.12)
- úloha (12.14)
- všechny úlohy ze Cvičení 12.4
- úloha LP vzniklá ve Cvičení 12.9
- Příklad 12.10.

(*) Dále pro každou úlohu zkuste interpretovat větu o silné dualitě (tj. úvahou odvod'te, jaká je optimální hodnota duální úlohy, a tato musí být stejná jako optimální hodnota primární úlohy) a podmínky komplementarity, podobně jako v Příkladu 14.6. Interpretace vět o dualitě je obecně velmi netriviální, takže většinou se vám to nepodaří – ale aspoň to zkuste.

14.3. Dokažte bez užití algoritmu na řešení LP, že $\mathbf{x} = (1, 1, 1, 1)$ je optimální řešení úlohy

$$\begin{array}{l} \min \quad [47 \ 93 \ 17 \ -93] \mathbf{x} \\ \text{za podm.} \quad \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \mathbf{x} \leq \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix} \end{array}$$

14.4. Napište duální úlohu k dopravnímu problému (12.8).

$$\text{Řešení: } \max \left\{ \sum_{i=1}^m a_i p_i + \sum_{j=1}^n b_j q_j \mid p_i \in \mathbb{R}, q_j \in \mathbb{R}, p_i + q_j \leq c_{ij} \right\}$$

Kapitola 15

Konvexní optimalizační úlohy

15.1 Třídy optimalizačních úloh

Optimalizační úlohy ve tvaru (11.7) se taxonomizují podle druhu funkcí f, g_i, h_i . Pro každou třídu existují specializované algoritmy schopné najít lokální minimum¹.

Lineární programování (LP)

V *lineárním programování* jsou všechny funkce f, g_i, h_i afinní. Jde tedy v jistém smyslu o nej-jednoduší případ konvexní optimalizační úlohy. Přesto jsme viděli v Kapitole 12, že již tento jednoduchý případ má velmi mnoho aplikací.

Kvadratické programování (QP)

V *kvadratickém programování* jsou funkce g_i, h_i afinní a funkce f je kvadratická konvexní, tedy $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$, kde matice \mathbf{A} je pozitivně semidefinitní.

Příklad 15.1. Při řešení soustavy ve smyslu nejmenších čtverců počítáme konvexní QP bez omezení $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A} \mathbf{x} - \mathbf{b}\|_2^2$.

Tuto úlohu lze všelijak modifikovat, např. můžeme přidat omezení $\mathbf{c} \leq \mathbf{x} \leq \mathbf{d}$, tj. každá proměnná x_j musí být v intervalu $[c_j, d_j]$. To vede na konvexní QP s omezeními. \square

Příklad 15.2. Hledání řešení lineární soustavy s nejmenší normou vede na úlohu $\min\{\mathbf{x}^T \mathbf{x} \mid \mathbf{A} \mathbf{x} = \mathbf{b}\}$, což je konvexní QP s omezeními. \square

Příklad 15.3. Chceme spočítat vzdálenost polyedrů

$$P_1 = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1\}, \quad P_2 = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}_2 \mathbf{x} \leq \mathbf{b}_2\}$$

danou jako $d(P_1, P_2) = \inf\{\|\mathbf{x}_1 - \mathbf{x}_2\|_2 \mid \mathbf{x}_1 \in P_1, \mathbf{x}_2 \in P_2\}$. Úloha vede na QP

$$\min\{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \mid \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n, \mathbf{A}_1 \mathbf{x}_1 \leq \mathbf{b}_1, \mathbf{A}_2 \mathbf{x}_2 \leq \mathbf{b}_2\}.$$

Pokud se polyedry protínají, jejich vzdálenost je nula. Pokud je aspoň jeden polyedr prázdný, úloha je nepřipustná. \square

¹ Viz např. <http://www.neos-guide.org>.

Kvadratické programování s kvadratickými omezeními (QCQP)

Obecnější variantou je *kvadratické programování s kvadratickými omezeními* (QCQP, *quadratically constrained quadratic programming*), kde všechny funkce f, g_i, h_i jsou kvadratické. Úloha je konvexní jen tehdy, když kvadratické funkce f, g_i jsou konvexní (tj. s pozitivně semidefinitní maticí) a funkce h_i jsou afinní.

Příklad 15.4. Na závěr ukažme jednoduchou konvexní úlohu, která na první pohled nespadá do žádné z uvedených tříd. Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ a chceme minimalizovat funkci

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{a}_i - \mathbf{x}\|_2 \quad (15.1)$$

přes $\mathbf{x} \in \mathbb{R}^n$. Řešení této úlohy je známo jako *geometrický medián*. Pro $n = 1$ se funkce redukuje na $f(x) = \sum_{i=1}^m |x - a_i|$, jejímž minimem je obyčejný medián.

Pro případ $n = 2$ má úloha jednoduchý mechanický model. Do vodorovného prkna vyvrtáme díry o souřadnicích \mathbf{a}_i . Každou dírou provlečeme provázek. Provázky jsou nahoře svázané uzlem do jednoho bodu a dole mají závaží o stejné hmotnosti. Poloha uzlu je \mathbf{x} . Hodnota $f(\mathbf{x})$ je potenciální energie soustavy a ustálený stav odpovídá minimu $f(\mathbf{x})$. \square

Přidat konvexní úlohu na analytický střed polyedru z Boyda: $\max\{\sum_i \log(\mathbf{a}_i^T \mathbf{x} - b_i) \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$.

15.2 Příklady nekonvexních úloh

Příklad 15.5. Řešení homogenní lineární soustavy ve smyslu nejmenších čtverců vede na úlohu

$$\min\{\|\mathbf{A}\mathbf{x}\|_2^2 \mid \mathbf{x}^T \mathbf{x} = 1\}. \quad (15.2)$$

To je instance QCQP, ale není to konvexní úloha kvůli omezení $\mathbf{x}^T \mathbf{x} = 1$. Dokonce ani nejde na konvexní úlohu transformovat. Je jasné, že množina $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{x} = 1\}$ není konvexní. Někdo by si mohl myslet, že omezení $\mathbf{x}^T \mathbf{x} = 1$ lze nahradit konvexním omezením $\mathbf{x}^T \mathbf{x} \leq 1$, podobně jako ve Cvičení 12.6. To ale nelze, neboť

$$\min\{\|\mathbf{A}\mathbf{x}\|_2^2 \mid \mathbf{x}^T \mathbf{x} = 1\} \neq \min\{\|\mathbf{A}\mathbf{x}\|_2^2 \mid \mathbf{x}^T \mathbf{x} \leq 1\} = 0.$$

My ale víme, že úlohu (15.2) lze řešit pomocí SVD, protože hledáme nadrovinu s normálovým vektorem \mathbf{x} , která minimalizuje součet čtverců kolmých vzdáleností řádků $\mathbf{a}_1, \dots, \mathbf{a}_m$ matice \mathbf{A} k nadrovině. \square

Tento příklad změnit na SVD.

V tomto příkladě měla nekonvexní úloha jediné lokální minimum. To je ale řídká výjimka – v naprosté většině mají nekonvexní úlohy mnoho lokálních extrémů.

Příklad 15.6. Uveďme příklad, na kterém bude na první pohled vidět, že nekonvexní úloha může mít velmi mnoho lokálních minim. Řešme úlohu

$$\min\{-\mathbf{x}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}. \quad (15.3)$$

Množina přípustných řešení je hyperkrychle, $X = [-1, 1]^n$. Účelová funkce $f(\mathbf{x}) = -\mathbf{x}^T \mathbf{x}$ je konkávní. Je očividné, že funkce f má na množině X lokální minimum v každém vrcholu hyperkrychle X (nakreslete si obrázek pro $n = 2$, tedy pro obyčejný čtverec!). Pro n proměnných má úloha 2^n lokálních minim. Připomeňme, že konvexní polyedr popsáný polynomiálním počtem lineárních nerovnic může mít exponenciální počet vrcholů (viz §12.1.2).

V tomto případě jsou lokální minima všechna stejná, tedy úlohu snadno vyřešíme. Ale již mírnou modifikací úlohy se stane nalezení globálního optima prakticky nemožné. Uvažujme úlohu

$$\min\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}\}. \quad (15.4)$$

Je jasné, že pro $\mathbf{A} = -\mathbf{I}$ dostaneme úlohu (15.3). Je známo, že neexistuje algoritmus, který by pro libovolnou (tedy ne nutně pozitivně semidefinitní) matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ vyřešil úlohu (15.4) v čase, který je shora omezen polynomiální funkcí čísla n . \square

Uvedme dále praktičtější příklady.

Příklad 15.7. Mějme m bodů v rovině $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^2$. Úkolem je rozmístit dalších n bodů $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^2$ tak, aby nejdelší vzdálenost bodu \mathbf{a}_i k nejbližšímu bodu \mathbf{x}_j byla nejmenší. Tedy minimalizujeme účelovou funkci

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \max_{i=1}^m \min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\| \quad (15.5)$$

přes vektory $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^2$. Máme $f: \mathbb{R}^{2n} \rightarrow \mathbb{R}$, tedy přesněji můžeme říci, že minimalizujeme funkci f přes jediný vektor $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{2n}$.

Úloha je známá jako *shlukování*. Jako motivaci si představme optimální rozmístění cisteren ve vesnici, kde občas neteče voda. Zde \mathbf{a}_i jsou souřadnice domů a \mathbf{x}_j jsou souřadnice cisteren. Chceme, aby obyvatel každého domu měl k nejbližší cisterně co nejbližší.

Je funkce f konvexní? Funkce $g_i(\mathbf{x}_1, \dots, \mathbf{x}_n) = \|\mathbf{a}_i - \mathbf{x}_j\|$ jsou konvexní pro každé i . Ale funkce $h_i(\mathbf{x}_1, \dots, \mathbf{x}_n) = \min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\|$ již konvexní být nemusí (Větu 11.7 nelze použít, ta hovoří o *maximu* konvexních funkcí). Tedy ani funkce $f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \max_{i=1}^m h_i(\mathbf{x}_1, \dots, \mathbf{x}_n)$ nemusí být konvexní. Dobře si ujasněte význam funkcí g_i, h_i, f a jejich definiční obory!

Tím, že se nám nepodařilo dokázat konvexitu funkce f , jsme samozřejmě nedokázali její nekonvexitu. A už vůbec jsme nedokázali, že funkce f má více než jedno lokální minimum. Nicméně je známo, že neexistuje algoritmus, který by našel optimální řešení úlohy (15.5) pro libovolný soubor bodů \mathbf{a}_i v čase, který je polynomiální funkcí čísel m a n . V praktické situaci tedy nezbyvá nic jiného, než použít algoritmus, který najde pouze přibližné optimum. Takovým algoritmem je např. *k-means*.

Úlohu (15.5) lze modifikovat nahrazením maxima součtem,

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^m \min_{j=1}^n \|\mathbf{a}_i - \mathbf{x}_j\|. \quad (15.6)$$

Opět se jedná o nekonvexní úlohu. Jaký význam má tato formulace? \square

15.2.1 Celočíslné programování

Významnou skupinou nekonvexních úloh jsou úlohy, ve kterých přípustná řešení nabývají pouze celočíselných hodnot. Z nich nejvýznamější je **celočíslné lineární programování** (ILP, *integer linear programming*)

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^n, \mathbf{A} \mathbf{x} \geq \mathbf{b}\}. \quad (15.7)$$

Rozdíl oproti obyčejnému LP je v tom, že místo $\mathbf{x} \in \mathbb{R}^n$ je $\mathbf{x} \in \mathbb{Z}^n$. Často proměnné nabývají dokonce pouze dvou stavů, tedy $\mathbf{x} \in \{0, 1\}^n$, pak mluvíme o *binární LP* nebo *0-1 LP*. Množina přípustných řešení této úlohy je nekonvexní, obsahuje konečný počet izolovaných bodů.

Množinu přípustných řešení můžeme napsat dvěma způsoby:

$$X = \{ \mathbf{x} \in \mathbb{Z}^n \mid \mathbf{Ax} \geq \mathbf{b} \} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b} \} \cap \mathbb{Z}^n.$$

Druhý způsob říká, že X jsou body celočíselné mřížky \mathbb{Z}^n ležící uvnitř konvexního polyedru $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b} \}$. Tento polyedr je množinou přípustných řešení obyčejného LP.

Zatímco LP je snadné vyřešit (LP je řešitelné v polynomiálním čase), ILP je NP-úplné. Mnoho úloh kombinatorické optimalizace (např. úloh na grafech) se pohodlně formuluje jako ILP.

Příklad 15.8. V úloze o pokrytí množiny máme dán systém množin $F = \{S_1, \dots, S_n\}$ (tedy S_i jsou množiny a F je množina množin) a úkolem je vybrat z něj co nejmenší podmnožinu, která má stejné sjednocení jako původní systém. Jedná se o jednu z klasických NP-úplných úloh.

Formulujme ji jako ILP. Proměnné budou $x_1, \dots, x_n \in \{0, 1\}$, kde $x_i = 1$ indikuje, že množina S_i byla vybrána.

$$\begin{aligned} & \min \sum_{i=1}^n x_i \\ & \text{za podmínek} \quad \sum_{i|e \in S_i} x_i \geq 1, \quad \forall e \in S_1 \cup \dots \cup S_n \\ & \quad \quad \quad x_1, \dots, x_n \in \{0, 1\} \end{aligned}$$

Necht' např. $F = \{\{a, b\}, \{b, c\}, \{a, c\}\}$. Existují tři optimální pokrytí, každé obsahuje dvě z daných tří množin: $\mathbf{x} = (1, 1, 0)$, $\mathbf{x} = (1, 0, 1)$ a $\mathbf{x} = (0, 1, 1)$. Každé z nich má optimální hodnotu ILP rovnu 2. \square

Přidat: Někdy má ILP vždy celočíselné řešení. Příklad: assignment problem:

$$\min \left\{ \sum_i c_{if(i)} \mid f: \{1, \dots, n\} \rightarrow \{1, \dots, n\} \text{ bijekce} \right\}$$

(kolik je bijekcí?) což se dá psát jako ILP

$$\min \left\{ \sum_{i,j} c_{ij} x_{ij} \mid \sum_j x_{ij} = 1, \sum_i x_{ij} = 1, x_{ij} \in \{0, 1\} \right\}$$

15.2.2 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

$$Y \supset X \implies \min_{\mathbf{x} \in Y} f(\mathbf{x}) \leq \min_{\mathbf{x} \in X} f(\mathbf{x}). \quad (15.8)$$

Jak toho lze použít? Mějme úlohu $\min_{\mathbf{x} \in X} f(\mathbf{x})$, kterou je obtížné vyřešit. Dejme tomu, že obtížnost pramení ze složitosti množiny přípustných řešení X . Nahradíme množinu X 'jednodušší' množinou $Y \supset X$ a řešíme úlohu $\min_{\mathbf{x} \in Y} f(\mathbf{x})$. Když budeme mít štěstí, bude nerovnost v (15.8) platit s rovností, tedy obě optima budou stejná. Když ne, získáme alespoň dolní mez na optimální řešení, které může být různými způsoby užitečné.

Typická situace je, že množina X je nekonvexní a my ji nahradíme vhodnou konvexní množinou $Y \supset X$. Pokud funkce f je konvexní, získáme konvexní úlohu. Mluvíme o **konvexní relaxaci**. Tak například konvexní relaxace úlohy ILP (15.7) je úloha LP

$$\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} \geq \mathbf{b}\}. \quad (15.9)$$

Zde máme $X = \mathbb{Z}^n$, $Y = \mathbb{R}^n \supset X$, a $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$.

Příklad 15.9. V úloze 15.8 je optimální hodnota LP relaxace rovna $\frac{3}{2}$, odpovídající $\mathbf{x} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. □

Přidat příklady na alternující minimalizaci: shlukování, low-rank approximation, non-negative matrix factorization $\min_{X, Y \geq 0} \|\mathbf{A} - \mathbf{X}\mathbf{Y}\|_F$.

15.3 Cvičení

15.1. Najděte explicitní řešení pro následující úlohy QCQP (\mathbf{A}, \mathbf{B} jsou pozitivně definitní):

a) $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A}\mathbf{x} \leq 1\}$

Nápověda: Viz Cvičení 12.6.

b) $\min\{\mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, (\mathbf{x} - \mathbf{b})^T \mathbf{A}(\mathbf{x} - \mathbf{b}) \leq 1\}$

Nápověda: substituuje $\mathbf{y} = \mathbf{x} - \mathbf{b}$.

c) $\min\{\mathbf{x}^T \mathbf{B}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^T \mathbf{A}\mathbf{x} \leq 1\}$

15.2. Formulujte úlohu $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_4$ jako konvexní QCQP.

15.3. Máme konvexní funkci jedné proměnné $f: \mathbb{R} \rightarrow \mathbb{R}$. Dáme do grafu funkce žebřík o délce 1 tak, aby oba konce ležely na grafu. Předpokládáme-li, že tření mezi žebříkem a grafem je nulové, zaujme žebřík stav lokálního minima potenciální energie (která je přímo úměrná výšce středu žebříku). Zformulujte jako optimalizační úlohu. Bude tato úloha konvexní? Pokud ne, najděte situaci, kdy potenciální energie bude mít více než jedno lokální minimum.

15.4. Dokažte, že účelové funkce vystupující v následujících úlohách jsou nekonvexní:

a) úloha (15.5)

b) úloha (15.6)

c) Příklad 10.6

d) Cvičení 10.2

Možný postup je metoda Monte Carlo: V Matlabu generujte náhodně (příkazem `randn`) potřebné vektory a čísla tak dlouho, dokud neporuší podmínku (11.3).

Kapitola 16

Vícekriteriální optimalizace

16.1 Uspořádání na množině

Binární relace na množině X je množina $R \subseteq X \times X$. Binární relace je

- **reflexivní**, když $(x, x) \in R$ pro každé $x \in X$,
- **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 $x = y$.

Definice 16.1. **Kvasi-uspořádání** na množině X je binární relace na X , která je reflexivní a transitivní. **Částečné uspořádání** (krátce jen **uspořádání**) na množině X je binární relace na X , která je reflexivní, transitivní a antisymetrická.

Relace kvasi-uspořádání a uspořádání se obvykle značí infixově symbolem \preceq , tedy místo $(x, y) \in R$ píšeme $x \preceq y$.

Prvky $x, y \in X$ jsou **srovnatelné**, když $x \preceq y$ nebo $y \preceq x$ nebo obojí. (Kvasi-)uspořádání je **úplné** (neboli **totální**), když každé dva prvky z X jsou srovnatelné.

Příklad 16.1. Příklady uspořádání:

- $X = \mathbb{R}$ a \preceq je přirozené uspořádání reálných čísel. Toto uspořádání je úplné.
- $X \subseteq 2^U$ 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.
- $X = \mathbb{N}$ a \preceq je relace dělitelnosti, tj. $x \preceq y$ právě když x dělí y . □

Nejvíce nás ovšem zajímá případ, kdy $X = \mathbb{R}^n$.

Příklad 16.2. Příklady uspořádání na množině \mathbb{R}^n :

- Uspořádání 'po složkách': $\mathbf{x} \leq^n \mathbf{y}$ právě když $x_i = y_i$ pro všechna $i = 1, \dots, n$.
- Lexikografické uspořádání definované jako

$$\mathbf{x} \leq_{\text{lex}} \mathbf{y} \iff (\mathbf{x} = \mathbf{y}) \text{ nebo } (\exists m \in \{1, \dots, n\})(\forall i < m)(x_i = y_i, x_m < y_m).$$

Toto uspořádání je úplné. □

Příklad 16.3. Příklady kvasi-uspořádání na množině \mathbb{R}^n :

- $\mathbf{x} \preceq \mathbf{y}$ právě když $x_i = y_i$ pro všechna $i = 1, \dots, n$ (toto kvazi-uspořádání je symetrická relace, tedy ekvivalence).
- $\mathbf{x} \preceq \mathbf{y}$ právě když $\sum_{i=1}^n x_i \leq \sum_{i=1}^n y_i$.
- Max-uspořádání: $\mathbf{x} \preceq \mathbf{y}$ právě když $\max_{i=1}^n x_i = \max_{i=1}^n y_i$
- Leximax uspořádání: $\mathbf{x} \preceq \mathbf{y}$ právě když $\text{sort}(\mathbf{x}) \leq_{\text{lex}} \text{sort}(\mathbf{y})$, kde zobrazení $\text{sort}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ seřadí složky vektoru sestupně, např. $\text{sort}(0, 3, 2, 8) = (8, 3, 2, 0)$. \square

Definice 16.2. Prvek $x \in X$ se nazývá (vzhledem ke (kvazi-)uspořádání \preceq)

- **minimální prvek** množiny X , když $y \preceq x$ implikuje $x \preceq y$, pro všechna $y \in X$.
- **nejmenší prvek** množiny X , když $x \preceq y$, pro všechna $y \in X$.

Pro totální (kvazi-)uspořádání oba pojmy splývají.

16.2 Úlohy vícekritériální optimalizace

V klasické optimalizaci jsme se zabývali úlohami typu

$$\min_{x \in X} f(x), \quad (16.1)$$

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 (kvazi-)uspořádání na množině Y . Pak úlohou (16.1) 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 16.4. 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 (kvazi-)uspořádání na množině \mathbb{R}^2 použít pro rozhodování.

Rozhodujme se vzhledem k uspořádání 'po složkách' \leq^2 . Vzhledem k tomuto uspořádání nemá množina $\mathbf{f}(X)$ nejmenší prvek, neboli kritéria jsou v konfliktu. Její minimální prvky jsou $\mathbf{f}(\text{Opel Astra}) = (15, 7.0)$ a $\mathbf{f}(\text{Ford Focus}) = (14, 7.5)$.

Rozhodujme se vzhledem k lexikografickému uspořádání, přesněji nejprve se rozhodujme dle ceny a pak dle spotřeby. Nyní minimální prvek množiny $\mathbf{f}(X)$ je $\mathbf{f}(\text{Ford Focus}) = (14, 7.5)$. \square

¹ Angl. *multicriteria optimization*. Názvosloví ovšem není jednotné, jindy se používají názvy *multiobjective optimization* nebo *vector optimization* (neboť hodnoty zobrazení f jsou vektory).

Příklad 16.5. Chceme řešit přibližně přeuračenu nehomogenní soustavu $\mathbf{Ax} \approx \mathbf{b}$, ale zároveň chceme, aby velikost vektoru \mathbf{x} byla malá. Řešíme tedy úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^n} (\|\mathbf{Ax} - \mathbf{b}\|_2, \|\mathbf{x}\|_2).$$

Jaké jsou minimální prvky množiny $\mathbf{f}(\mathbb{R}^2)$ vzhledem k uspořádání po složkách? Je jich nekonečně mnoho, obrázek. \square

Příklad 16.6. 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 $\mathbf{a}_1, \dots, \mathbf{a}_n$ (toto tvrzení podrobně nedokazujeme). 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. \square

Dále: Více o Pareto-optimálních řešeních, metoda nezáporné kombinace (skalarizace), předvést na regularizaci.