Optimalizace

Zárodek skript k předmětu A4B33OPT.
Text je neúplný a v průběhu semestru je doplňován a vylepšován.
Toto je verze ze dne **13. prosince 2011**.

Tomáš Werner

České vysoké učení technické
Fakulta elektrotechnická

Obsah

Kapitola 1

Úvod

1.1 Disciplína optimalizace

Optimalizace (přesněji matematická optimalizace) se zabývá minimalizací (či maximalizací) funkcí mnoha proměnných za případných omezujících podmínek. Tato formulace pokrývá mnoho úloh z inženýrské praxe i přírodních věd: často přeci chceme něco udělat 'nejlépe' v rámci 'daných možností'. Umět rozpoznávat optimalizační problémy kolem sebe je inženýrovi velmi užitečné. Optimalizace, též zvaná *matematické programování*, je část aplikované matematiky, ležící na pomezí matematické analýzy, lineární algebry a informatiky. Je to moderní obor, který se rychle rozvíjí.

Příklady problémů, které vedou na optimalizační úlohy:

- Aproximuj naměřenou funkční závislost funkcí z dané třídy funkcí.
- Investuj 1000 Kč do daných druhů akcií tak, aby očekávaný výnos byl velký a riziko malé.
- Rozmísti daný počet prodejen po městě tak, aby každý člověk měl do prodejny blízko.
- Najdi průběh řídícího signálu ruky robota tak, aby se dostala z místa A do místa B po dráze minimální délky (příp. minimálního času či výdaje energie) a bez kolize.
- Reguluj přívod plynu do kotle tak, aby teplota v domě byla blízká kýžené teplotě.
- Navrhni plošný spoj daného zapojení, aby délka spojů byla nejmenší.
- Najdi nejkratší cestu v počítačové síti.
- Vyhledej nejlepší spojení v jízdním řádu z místa A do místa B.
- Navrhni nejlepší školní rozvrh.
- Postav most o dané nosnosti při nejmenší spotřebě materiálu.
- Nauč umělou neuronovou síť.

Mimo inženýrskou praxi je optimalizace významná v přírodních vědách. Většinu fyzikálních zákonů lze formulovat tak, že nějaká veličina nabývá extrémální hodnoty. Živé organismy v každém okamžiku řeší, vědomě či podvědomě, množství optimalizačních úloh – např. se rozhodují pro nejlepší z možných chování.

V tomto kursu se nenaučíte řešit všechny tyto úlohy – už proto, že některé jsou velmi těžké. Ale naučíte se rozpoznat druh a obtížnost úloh a dostanete základy pro řešení těch snadnějších a přibližné řešení těch obtížnějších. Spektrum úloh, které dokážete řešit, se ještě podstatně rozšíří po absolvování navazujícího kursu *Kombinatorická optimalizace*.

1.2 Matematické značení

Množiny

$\{a_1, \ldots, a_n\}$	množina s prvky $a_1, \ldots, a_n$
$a \in A$	prvek a patří do množiny A (neboli a je prvkem A)
$A \subseteq B$	množina A je podmnožinou množiny B (každý prvek z X patří do Y)
$A = B$	množina A je rovna množině B, platí zároveň $A \subseteq B$ a $B \subseteq A$
$\{\, a \in A \mid \varphi(a) \,\}$	množina prvků z A s vlastností φ. Někdy zkracujeme na $\{\, a \mid \varphi(a) \,\}$.
$A \cup B$	sjednocení množin, množina $\{\, a \mid a \in A$ nebo $a \in B \,\}$
$A \cap B$	průnik množin, množina $\{\, a \mid a \in A$ a zároveň $a \in B \,\}$
$(a_1, \ldots, a_n)$	uspořádaná n-tice prvků $a_1, \ldots, a_n$
$A \times B$	kartézský součin množin, množina $\{\, (a, b) \mid a \in A,\ b \in B \,\}$
A^n	kartézský součin n stejných množin, $A^n = A \times \cdots \times A$ (n-krát).

Zobrazení

Zápisem $f\colon A \to B$ rozumíme zobrazení z množiny A (zvané **definiční obor**) do množiny B. Formální definice je tato: podmnožina f kartézského součinu $A \times B$ (tedy *relace*) se nazývá zobrazení, platí-li $(a, b) \in f,\ (a, b') \in f \;\Rightarrow\; b = b'$. Neformálně si představujeme zobrazení jako černou skříňku (např. počítačovou funkci), která přiřadí každému prvku $a \in A$ jediný prvek $b = f(a) \in B$. Přísně vzato, 'zobrazení' (*mapping, map*) znamená přesně to samé jako 'funkce' (*function*), ovšem slovo 'funkce' se obvykle používá pouze pro zobrazení do číselných množin (tedy $B = \mathbb{R}$, $\mathbb{Z}$, $\mathbb{C}$ apod.).

Pro množinu obrazů všech vzorů s vlastností φ se používá zkratka $\{\, f(a) \mid a \in A,\ \varphi(a) \,\} = \{\, b \in B \mid b = f(a),\ a \in A,\ \varphi(a) \,\}$ nebo pouze $\{\, f(a) \mid \varphi(a) \,\}$, pokud je A jasné z kontextu. Např. množina $\{\, x^2 \mid -1 < x < 1 \,\}$ je polouzavřený interval $\langle 0, 1)$. Obraz množiny A v zobrazení f budeme často značit $f(A) = \{\, f(a) \mid a \in A \,\}$.

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 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.

Číselné množiny

$\mathbb{N}$	množina přirozených čísel
$\mathbb{Z}$	množina celých čísel
$\mathbb{Q}$	množina racionálních čísel
$\mathbb{R}$	množina reálných čísel
$\mathbb{R}_+$	množina nezáporných reálných čísel
$\mathbb{R}_{++}$	množina kladných reálných čísel
$\langle x_1, x_2)$	polouzavřený interval reálných čísel, $\{\, x \in \mathbb{R} \mid x_1 \le x < x_2 \,\}$
$\mathbb{C}$	množina komplexních čísel

Vektory a matice

$\mathbb{R}^{m\times n}$	množina reálných matic rozměru $m \times n$ (tedy s m řádky a n sloupci)
$\mathbb{R}^n$	množina sloupcových vektorů, ztotožněná s množinou $\mathbb{R}^{n\times 1}$ jednosloupcových matic
$\mathbb{R}^{1\times n}$	množina řádkových vektorů
$\mathbf{x}\cdot\mathbf{y}$	skalární součin vektorů
$\|\mathbf{x}\|_p$	p-norma vektoru $\mathbf{x}$
$\|\mathbf{x}\|$	norma vektoru $\mathbf{x}$, bez dalšího upřesnění eukleidovská norma $\|\mathbf{x}\|_2$

Vektory a matice značíme **tučně**, skaláry *kurzívou.*

Symbol $f\colon \mathbb{R}^n \to \mathbb{R}$ značí funkci n proměnných, která vektoru $\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^n$ přiřadí skalár $f(\mathbf{x}) = f(x_1,\dots,x_n) \in \mathbb{R}$. Jméno funkce píšeme kurzívou, neboť její hodnoty jsou skaláry.

Symbol $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ značí zobrazení, které vektoru $\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^n$ přiřadí vektor $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1,\dots,x_n) = (f_1(\mathbf{x}),\dots,f_m(\mathbf{x})) = (f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n)) \in \mathbb{R}^m$. Jméno zobrazení píšeme tučně, neboť jeho hodnoty jsou vektory. Funkce $f_1,\dots,f_m\colon \mathbb{R}^n \to \mathbb{R}$ jsou složky zobrazení $\mathbf{f}$.

Extrém funkce na množině

Pro minimum funkce $f\colon X \to \mathbb{R}$ na množině X se užívá značení

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

Pro množinu prvků X, ve kterých se minimum nabývá, se používá symbol **argument minima**:

$$\operatorname*{argmin}_{x\in X} f(x) = \{\, x^* \in X \mid f(x^*) = \min_{x\in X} f(x) \,\}.$$

Pro maximum je to analogické. Minima a maxima se souhrnně nazývají **extrémy** nebo **optima**.

Příklad 1.1.

1. $\min_{x\in\mathbb{R}} |x-1| = \min\{\, |x-1| \mid x \in \mathbb{R} \,\} = 0$, $\operatorname*{argmin}_{x\in\mathbb{R}} |x-1| = \{1\}$
2. Nechť $(a_1,a_2,\dots,a_5) = (1,2,3,2,3)$. Pak $\max_{i=1}^{5} a_i = 3$, $\operatorname*{argmax}_{i=1}^{5} a_i = \{3,5\}$. □

1.3 Formulace optimalizačních úloh

Optimalizační úlohy lze formulovat jako hledání minima dané reálné funkce na dané množině:

$$\min_{x\in X} f(x). \tag{1.1}$$

V optimalizaci se užívá následující názvosloví. Funkce f se nazývá **účelová** (také pokutová, cenová, kriteriální) funkce. Množina X se nazývá množina **přípustných řešení** – což je vlastně protimluv, protože prvky X nejsou řešeními úlohy. Prvkům $x^* \in X$ pro které $f(x^*) = \min_{x\in X} f(x)$ se pak říká **optimální řešení**. Optimální řešení může být jedno, více, nebo nemusí existovat.

Formulace (1.1) je velmi obecná a abstraktní, 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 vektory, mluvíme o *spojité optimalizaci*. Příkladem je úloha lineárního programování.
- Pokud množina X obsahuje funkce, mluvíme o *variačním počtu*. Příkladem je nalézt tvar rovinné křivky, která při dané délce obepíná co největší obsah.

Tento kurs se věnuje hlavně spojité optimalizaci. Zde prvky množiny X jsou n-tice reálných proměnných $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, tedy $X \subseteq \mathbb{R}^n$. Zde zapisujeme $\mathbf{x}$ tučně, protože jej považujeme za vektor. Množina přípustných řešení X je definována jako množina všech $\mathbf{x} \in \mathbb{R}^n$ splňujících rovnice

$$\begin{aligned} g_i(\mathbf{x}) &\leq 0, \quad i = 1, \ldots, m \\ h_i(\mathbf{x}) &= 0, \quad i = 1, \ldots, \ell \end{aligned} \tag{1.2}$$

pro dané reálné funkce n proměnných $g_1, \ldots, g_m$ a $h_1, \ldots, h_\ell$. Úloha se zapisuje také jako

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, \ell. \end{aligned} \tag{1.3}$$

Příklad 1.2. Máme najít bod v rovině, který leží na kružnici s jednotkovým poloměrem a se středem v počátku a který je nejblíže danému bodu $\mathbf{a}$. Zde máme $n = 2$, $m = 0$, $\ell = 1$, $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$, $h_1(\mathbf{x}) = \|\mathbf{x}\| - 1$. Řešíme úlohu $\min\{\, \|\mathbf{a} - \mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^2,\ \|\mathbf{x}\| = 1 \,\}$, tedy minimalizujeme $\|\mathbf{a} - \mathbf{x}\|$ za podmínky $\|\mathbf{x}\| = 1$. $\square$

Rovnice a nerovnice (1.2) jsou **omezující podmínky**, krátce **omezení**. Omezení tvaru $g_i(\mathbf{x}) \leq 0$ příp. $h_i(\mathbf{x}) = 0$ se nazývá omezení **typu nerovnosti** příp. **typu rovnosti**. Všimněte si, že omezení $h_i(\mathbf{x}) = 0$ je ekvivalentní dvěma omezením $h_i(\mathbf{x}) \leq 0$ a $-h_i(\mathbf{x}) \leq 0$, proto jsme vlastně omezení typu rovnosti v problému (1.3) nemuseli explicitně uvádět. Omezení mohou chybět ($m = \ell = 0$).

V závislosti na účelové funkci a omezeních může být nalezení optimálního řešení snadné, obtížné, nebo prakticky nemožné.

V tomto kursu se budeme zabývat těmito speciálními případy:

- $m = \ell = 0$, funkce f spojitá diferencovatelná.
  Vícerozměrná optimalizace bez omezení. Speciálně f kvadratická pozitivně definitní: lineární nejmenší čtverce.
- $m = 0$, $\ell > 0$, funkce f, h_i diferencovatelné.
  Zde se naučíte metodu Lagrangeových multiplikátorů.
- $m, \ell > 0$, funkce f, g_i, h_i direferencovatelné.
  Zobecnění metody Lagrangeových multiplikátorů: Karush-Kuhn-Tucker (KKT) podmínky.
- $m, \ell > 0$, funkce f, g_i, h_i afinní.
  Lineární programování (LP) a simplexová metoda.
- $m, \ell > 0$, funkce f kvadratická pozitivně semidefinitní, g_i, h_i afinní.
  Kvadratické programování (QP).

- $m, \ell > 0$, funkce f, g_i konvexní, h_i afinní.
  Konvexní programování, existuje jen jedno lokální minimum. Budeme se zabývat vlastnostmi konvexních množin a funkcí.

1.4 Cvičení

1.1. Najděte (úvahou) co nejjednodušší popis následujících množin:

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

1.2. Mějme množinu bodů v rovině $X = \langle -1, 1 \rangle \times \{0\} = \{\, (x, 0) \mid -1 \leq x \leq 1 \,\} \subseteq \mathbb{R}^2$. Načrtněte následující množiny:

a) $\left\{\, \mathbf{y} \in \mathbb{R}^2 \,\middle|\, \min_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\| \leq 1 \,\right\}$
b) $\left\{\, \mathbf{y} \in \mathbb{R}^2 \,\middle|\, \max_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\| \leq 2 \,\right\}$

1.3. Formulujte (již však neřešte) následující úlohy ve tvaru $\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X \,\}$.

a) Najdi dvojici bodů v rovině, z nichž jeden je uvnitř čtverce se středem v počátku a jednotkovou stranou a druhý je uvnitř kruhu se středem v bodě $(2, 2)$ a jednotkovým poloměrem, tak aby si tyto dva body byly nejblíže.
b) Najdi dvě přirozená čísla se součtem 7 a nejmenším součinem.
c) Nechť $\mathbf{A}$ je matice rozměru $m \times n$, kde $m < n$, a $\mathbf{b}$ je vektor délky m. Předpokládejte, že matice má hodnost m. Najdi řešení $\mathbf{x}$ soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ tak, že délka vektoru $\mathbf{x}$ je minimální.
d) Ve vesnici je n chalup o daných souřadnicích. Najdi souřadnice pošty tak, aby pošťák měl k nejvzdálenější (měřeno vzdušnou čarou) chalupě co nejblíže.
e) Ve vesnici je n chalup o daných souřadnicích. Najdi souřadnice m pump tak, aby vzdálenost (vzdušnou čarou) od libovolné chalupy k nejbližší pumpě byla minimální.

1.4. Vyřešte následující úlohy:

a) $\min\{\, x^2 + y^2 \mid x > 0,\ y > 0,\ xy \geq 1 \,\}$
b) $\min\{\, (x - 2)^2 + (y - 1)^2 \mid x^2 \leq 1,\ y^2 \leq 1 \,\}$
c) $\max\{\, \mathbf{x} \cdot \mathbf{y} \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x} \cdot \mathbf{x} \leq 1 \,\}$ pro daný vektor $\mathbf{y} \in \mathbb{R}^n$ ($\mathbf{x} \cdot \mathbf{y}$ značí skalární součin). Zkuste nejprve pro $n = 1$, pak pro $n = 2$, pak zobecněte na libovolné n.
d) Najdi rozměry krabice bez víka o jednotkovém objemu a nejmenším povrchu.
e) Hledá se n-tice čísel $x_1, \ldots, x_n \in \{-1, +1\}$ tak, že jejich součin je kladný a jejich součet minimální. Jako výsledek napište hodnotu tohoto minimálního součtu v závislosti na n.

Kapitola 2

Vektory a matice

Zde zopakujeme potřebné partie lineární algebry a přidáme některé nové, které budeme potřebovat později. Z lineární algebry, kterou znáte z nižších ročníků, se budeme věnovat pouze algebře reálných vektorů konečné dimenze a matic konečné dimenze. Nepostupujeme formálně metodou definice-věta-důkaz, většinu tvrzení předložíme jako fakta bez důkazu nebo důkaz pouze naznačíme.

2.1 Lineární prostor

Nechť

- V je neprázdná množina,
- $+$ je binární operace $V \times V \to V$,
- $\cdot$ je binární operace $\mathbb{R} \times V \to V$.

Nechť jsou splněny tyto podmínky (axiomy):

- $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$
- $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- Existuje $\mathbf{y} \in V$ tak, že pro každé $\mathbf{x} \in V$ je $\mathbf{y} + \mathbf{x} = \mathbf{x}$. Značíme $\mathbf{y} = \mathbf{0}$.
- Pro každé $\mathbf{x} \in V$ existuje $\mathbf{y} \in V$ tak, že $\mathbf{x} + \mathbf{y} = \mathbf{0}$. Značíme $\mathbf{y} = -\mathbf{x}$.
- $1 \cdot \mathbf{x} = \mathbf{x}$
- $\alpha \cdot (\beta \cdot \mathbf{x}) = (\alpha \cdot \beta) \cdot \mathbf{x}$
- $\alpha \cdot (\mathbf{x} + \mathbf{y}) = (\alpha \cdot \mathbf{x}) + (\alpha \cdot \mathbf{y})$
- $(\alpha + \beta) \cdot \mathbf{x} = (\alpha \cdot \mathbf{x}) + (\beta \cdot \mathbf{x})$

Trojice $(V, +, \cdot)$ se pak nazývá **lineární prostor** (nebo **vektorový prostor**) nad tělesem reálných čísel. Prvkům množiny V se říká **vektory** a prvkům tělesa $\mathbb{R}$ **skaláry**. Místo $\alpha \cdot \mathbf{x}$ píšeme často jen $\alpha\mathbf{x}$.

Operace $+$ a $\cdot$ tělesa $\mathbb{R}$ pro jednoduchost značíme stejnými symboly jako operace lineárního prostoru. Kupř. $+$ někdy používáme jako operaci $V \times V \to V$ (tj. jako sčítání vektorů) a jindy jako operaci $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ (sčítání skalárů, prvků tělesa $\mathbb{R}$). Toto *přetěžování operátorů* významně zjednodušuje značení, ale student si ho musí být vědom a vždy si ujasnit ‘datové typy’ argumentů každého výskytu operátoru.

Příklad 2.1. Příklady lineárních prostorů:

1. Množina $\{\mathbf{0}\}$ (tedy množina obsahující pouze prvek $\mathbf{0}$) s operacemi definovanými jako $\mathbf{0}+\mathbf{0}=\mathbf{0}$ a $\alpha \cdot \mathbf{0}=\mathbf{0}$. Toto je nejmenší možný lineární prostor a říká se mu *triviální*.
2. Množina $\mathbb{R}^n$ všech n-tic reálných čísel $(x_1,\ldots,x_n)$, definujeme-li operaci $+$ jako sčítání po složkách a operaci $\cdot$ jako násobení všech složek skalárem.
3. Množina všech matic ('tabulek') reálných čísel velikosti $m \times n$, spolu s operacemi definovanými opět 'po složkách'.
4. Množina všech funkcí $f\colon X \to \mathbb{R}$, kde X je libovolná daná množina. Operace jsou přirozené sčítání funkcí a násobení skalárem 'v každém argumentu zvlášť'.
5. Množina všech polynomů $p\colon \mathbb{R} \to \mathbb{R}$ libovolného stupně. Operace opět dány přirozeně.

Množina všech polynomů stupně n není lineární prostor, protože součet dvou plynomů stupně n může být polynom stupně nižšího. □

Lineární kombinace vektorů $\mathbf{x}_1,\ldots,\mathbf{x}_k \in V$ je vektor

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k$$

pro nějaké skaláry $\alpha_1,\ldots,\alpha_k \in \mathbb{R}$. Lineární prostor je uzavřený vůči lineárním kombinacím, tj. lineární kombinace jakýchkoliv vektorů z V zůstává ve V. To snadno (indukcí) plyne z podmínek

- pro každé $\mathbf{x},\mathbf{y} \in V$ je $\mathbf{x}+\mathbf{y} \in V$,
- pro každé $\alpha \in \mathbb{R}$ a $\mathbf{x} \in V$ je $\alpha\mathbf{x} \in V$.

Rozumí se *samo sebou*, že tyto podmínky platí, neboť v definici lineárního prostoru jsme definovali operaci $+$ jako $V \times V \to V$ a operaci $\cdot$ jako $\mathbb{R} \times V \to V$.

Lineární podprostor (krátce **podprostor**) lineárního prostoru $(V,+,\cdot)$ je množina $U \subseteq V$, která je uzavřená vůči lineárním kombinacím (tj. každá lineární kombinace vektorů z U leží v U). Potom množina U s operacemi 'zděděnými' z původního prostoru V tvoří sama o sobě lineární prostor.

Vektory $\mathbf{x}_1,\ldots,\mathbf{x}_k$ jsou **lineárně závislé**, když existují skaláry $\alpha_1,\ldots,\alpha_k$ takové, že aspoň jeden z těchto skalárů je nenulový a

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}.$$

Ekvivalentně se dá říci, že libovolný vektor je lineární kombinací ostatních. Množina (konečná či nekonečná) $X \subseteq V$ je lineárně závislá, když nějaká její konečná podmnožina je lineárně závislá. V opačném případě je **lineárně nezávislá**.

Lineární obal množiny vektorů $X \subseteq V$ je množina

$$\operatorname{span} X = \{\, \alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k \mid \mathbf{x}_1,\ldots,\mathbf{x}_k \in X,\ \alpha_1,\ldots,\alpha_k \in \mathbb{R},\ k \in \mathbb{N} \,\}.$$

Tedy pokud je množina konečná, $X = \{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$, její lineární obal je množina všech lineárních kombinací vektorů $\mathbf{x}_1,\ldots,\mathbf{x}_k$. Pokud je X nekonečná, je její lineární obal množina všech lineárních kombinací všech konečných podmnožin X. Ekvivalentně lze říci, že lineární obal množiny $X \subseteq V$ je nejmenší lineární podprostor V obsahující X.

Báze lineárního prostoru V je lineárně nezávislá množina vektorů, jejíž lineární obal je celý prostor V. Každý lineární prostor má alespoň jednu bázi. Pokud lineární prostor má nějakou bázi s konečným počtem vektorů, pak jeho každá jiná báze má stejný počet vektorů. Tento počet je **dimenzí** lineárního prostoru, kterou značíme $\dim V$. Pokud V nemá bázi s konečným počtem prvků, je jeho dimenze nekonečná.

V našem kursu budeme potřebovat výhradně reálné lineární prostory s konečnou dimenzí. Základním příkladem takového prostoru je $\mathbb{R}^n$. Lze ukázat, že všechny reálné lineární prostory dimenze n jsou v jistém smyslu 'stejné' (isomorfní), tento příklad je tedy vlastně jediný. **Standardní bázi** prostoru $\mathbb{R}^n$ budeme značit

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0),\ \mathbf{e}_n = (0, \ldots, 0, 1).$$

Vektor ze samých jedniček budeme značit $\mathbf{1}_n = (1, \ldots, 1)$ nebo krátce $\mathbf{1}$.

2.2 Matice

Reálná **matice** rozměru $m \times n$ je zobrazení $\mathbf{A}$: $\{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{R}$. Hodnoty tohoto zobrazení značíme a_{ij}. Matici pomocí výčtu prvků zapisujeme tabulkou

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

Množinu všech reálných matic rozměru $m \times n$ (tj. s m řádky a n sloupci) značíme $\mathbb{R}^{m \times n}$.

Na množině $\mathbb{R}^{m \times n}$ zavedeme operace $+$ a $\cdot$ 'po prvcích'. Tím se tato množina stane lineárním prostorem dimenze mn. **Nulovou matici** (její všechny prvky jsou nuly) pak značíme $\mathbf{0}$.

Pro $m = n$ se matice nazývá **čtvercová**, pro $m \neq n$ **obdélníková**. Pro $m < n$ nazveme matici **širokou**, pro $m > n$ **úzkou** (toto názvosloví není standardní).

Čtvercová matice je **diagonální**, když má všude mimo hlavní diagonálu nuly, tedy $a_{ij} = 0$ pro $i \neq j$. **Jednotková** matice má na diagonále jedničky a mimo diagonálu nuly, značíme ji $\mathbf{I}$.

Transpozici matice $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ značíme $\mathbf{A}^T = [a_{ji}] \in \mathbb{R}^{n \times m}$. Platí $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$. Čtvercová matice se nazývá **symetrická**, když $\mathbf{A}^T = \mathbf{A}$, a **antisymetrická**, když $\mathbf{A}^T = -\mathbf{A}$.

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

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj},$$

což značíme $\mathbf{C} = \mathbf{AB}$. Vlastnosti maticového součinu:

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

Čtvercové matice obecně nekomutují, $\mathbf{AB} \neq \mathbf{BA}$.

Někdy je nutné sestavit matici z několika jejích podmatic (zvaných též *bloky*), což značíme

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}.$$

Rozměry jednotlivých bloků musí být kompatibilní. Všimněte si, že ve třetím příkladě jsou rozměry jednotkové matice $\mathbf{I}$ a nulové matice $\mathbf{0}$ určeny rozměry matic $\mathbf{A}$ a $\mathbf{D}$. Pro násobení matic sestavených z bloků platí podobné pravidlo jako pro matice sestavené ze skalárů:

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

2.2.1 Hodnost a inverze

Hodnost matice je maximální počet jejich lineárně nezávislých sloupců, což je totéž jako dimenze lineárního obalu sloupců matice. Hodnost značíme $\operatorname{rank} \mathbf{A}$. Platí $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{A}^T$, tedy místo pomocí sloupců jsme mohli hodnost definovat pomocí řádků. Obecně je vždy

$$\operatorname{rank} \mathbf{A} \leq \min\{m, n\}. \tag{2.1}$$

Pokud je $\operatorname{rank} \mathbf{A} = \min\{m, n\}$, říkáme, ža matice má **plnou hodnost**. Čtvercová matice s plnou hodností se nazývá **regulární**, jinak je **singulární**. Pro hodnost součinu matic platí

$$\operatorname{rank}(\mathbf{AB}) \leq \min\{\operatorname{rank} \mathbf{A}, \operatorname{rank} \mathbf{B}\}. \tag{2.2}$$

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

$$\mathbf{AB} = \mathbf{I}, \tag{2.3}$$

nazývá se matice $\mathbf{B}$ **pravá inverze** matice $\mathbf{A}$ a matice $\mathbf{A}$ **levá inverze** matice $\mathbf{B}$. Pravá či levá inverze nemusí existovat nebo nemusí být jediná. Např. pro $m < n$ rovnost (2.3) nikdy nenastane (proč?). Pravá inverze matice $\mathbf{A}$ existuje, právě když její řádky jsou lineárně nezávislé. Levá inverze matice $\mathbf{B}$ existuje, právě když její sloupce jsou lineárně nezávislé.

Pro $m = n$ (čtvercové matice) pravá inverze matice $\mathbf{A}$ existuje právě tehdy, když $\mathbf{A}$ je regulární (proto se regulární matici říká také **invertovatelná**). V tom případě je jediná a je rovna levé inverzi matice $\mathbf{A}$. Pak mluvíme pouze o **inverzi** matice $\mathbf{A}$ a značíme ji $\mathbf{A}^{-1}$. Máme tedy $\mathbf{AA}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$. Pro regulární matice $\mathbf{A}$ a $\mathbf{B}$ platí $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

2.2.2 Determinant

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

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

kde sčítáme přes všechny permutace n prvků $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, přičemž $\operatorname{sgn} \sigma$ označuje znaménko permutace. Některé jeho vlastnosti:

- Determinant je multilineární funkce sloupců matice, tj. je lineární funkcí libovolného sloupce, jsou-li všechny ostatní sloupce konstanty.
- 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

- $\det \mathbf{A}^T = \det \mathbf{A}$
- $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$
- $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$ (plyne z předchozího pro $\mathbf{B} = \mathbf{A}^{-1}$)

2.2.3 Ztotožnění vektorů a jednosloupcových matic

Široce se užívá následující konvence. Je jasné, že lineární prostor $\mathbb{R}^n$ všech n-tic $(x_1, \dots, x_n)$ a lineární prostor $\mathbb{R}^{n\times 1}$ všech matic s jediným sloupcem délky n jsou 'stejné' (isomorfní). Proto tyto prostory ztotožníme a budeme bez upozornění přecházet mezi oběma významy. Slovem 'vektor' v tomto smyslu rozumíme tedy objekt

$$\mathbf{x} = (x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T$$

Tedy zatímco obecně v lineární algebře slovo 'vektor' znamená prvek lineárního prostoru, v maticové algebře znamená jednosloupcovou matici. Totéž bychom samozřejmě mohli udělat s řádky (a někdo to tak i dělá, např. počítačoví grafikové). Proto matici $\mathbf{x} \in \mathbb{R}^{n\times 1}$ nazýváme přesněji **sloupcový vektor** a matici $\mathbf{x} \in \mathbb{R}^{1\times n}$ **řádkový vektor**.

Máme-li tedy matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ a vektor $\mathbf{x} \in \mathbb{R}^n$, výraz $\mathbf{y} = \mathbf{Ax} \in \mathbb{R}^n$ značí maticové násobení matice $m \times n$ maticí $n \times 1$, které definuje **násobení vektoru maticí**

$$y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Vektor $\mathbf{y}$ je lineární kombinace (s koeficienty x_j) sloupců matice $\mathbf{A}$.

Podobně, pro $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ je $\mathbf{x}^T\mathbf{y} = x_1y_1 + \cdots + x_ny_n$ je maticový součin řádkového vektoru $\mathbf{x}^T$ a sloupcového vektoru $\mathbf{y}$, jehož výsledkem je skalár. Naproti tomu $\mathbf{xy}^T$ je matice $m \times n$ (které se někdy říká *vnější součin* vektorů) hodnosti 1.

Mohli bychom si myslet, že když vektory lze považovat za matice rozměru $n \times 1$, tak skaláry lze považovat za matice rozměru 1×1. Není tomu tak: výraz $\alpha\mathbf{A}$ je syntakticky správně, ale není to maticový součin, protože vnitřní rozměr matic je různý. Pak navíc platí $\alpha\mathbf{A} = \mathbf{A}\alpha$.

2.3 Zločiny na maticích

Pro manipulaci s maticovými výrazy a rovnicemi je nutno získat podobnou zručnost, jakou máte pro manipulaci se skalárními výrazy a rovnicemi. K tomu je třeba nejenom znalost lineární algebry, ale i cvik – je nutno, aby vám 'prošlo rukama' co nejvíce výrazů a rovnic obsahujících matice a vektory. Při úpravách maticových výrazů dělají studenti často hrubé chyby, kterých se lze při alespoň minimální pozornosti vyvarovat. Výskyt takové chyby v testu či u zkoušky je *neomluvitelný*. Dále uvedeme typické kategorie těchto zločinů.

Výraz je nesmyslný kvůli rozměrům matic

První druh chyb je ten, že výraz nemá smysl kvůli rozměrům matic a vektorů. Rozlišíme dva druhy těchto chyb. Při **syntaktické chybě** pachatel napíše maticový výraz, který porušuje syntaktická pravidla. Příklady:

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

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

- Zcela odstrašující příklad je použití ‘zlomku’, např. $\dfrac{\mathbf{A}}{\mathbf{b}}$. To je možné jen tehdy, když ve jmenovateli je skalár.

Při **sémantické chybě** pachatel napíše výraz nebo udělá závěr, který sice neodporuje syntaxi, ale nedává smysl kvůli svému významu. Příklady:

- 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 $\operatorname{rank}\mathbf{A} = 5$, kde $\mathbf{A} \in \mathbb{R}^{3\times 5}$. Je chybné, protože každá pětice vektorů z $\mathbb{R}^3$ je lineárně závislá.

Příklad 2.2. Vidíme-li výraz $(\mathbf{A}^T\mathbf{B})^{-1}$, musíme okamžitě udělat 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 $m = k$. Výjimkou je případ, kdy $\mathbf{A}$ je skalár ($m = n = 1$) nebo $\mathbf{B}$ je skalár ($k = p = 1$) – pak by $\mathbf{A}^T\mathbf{B}$ byl sice zvláštní, ale syntakticky korektní zápis.
- 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.
- Pokud by matice $\mathbf{A}^T$ byla úzká nebo matice $\mathbf{B}$ široká, matice $\mathbf{A}^T\mathbf{B}$ by určitě 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, musí mít obě matice stejný rozměr a musí být čtvercové nebo úzké. Namítnete, že matice $\mathbf{A}^T\mathbf{B}$ nemusí mít inverzi ani za těchto podmínek – naším cílem však bylo pouze najít *nutné podmínky na rozměry matic*, aby výraz měl smysl. □

Použití neexistujících maticových identit

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

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

Neekvivalentní úpravy (ne)rovnic

Zde zločinec 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, ani když jsou obě matice čtvercové regulární – zkuste pro skaláry)

Další tipy pro práci s maticemi

- Pod maticové výrazy si malujte obdélníčky s rozměry matic, abyste měli přesnou 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.

2.4 Lineární zobrazení

Nechť U a V jsou lineární prostory. Zobrazení $\mathbf{f}\colon U \to V$ se nazývá **lineární**, pokud

- $\mathbf{f}(\mathbf{x}+\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{y})$
- $\mathbf{f}(\alpha\mathbf{x}) = \alpha\mathbf{f}(\mathbf{x})$

To je očividně (matematickou indukcí) ekvivalentní jediné podmínce

$$\mathbf{f}(\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k) = \alpha_1\mathbf{f}(x_1) + \cdots + \alpha_k\mathbf{f}(x_k), \tag{2.4}$$

tedy ‘zobrazení lineární kombinace je rovno lineární kombinaci zobrazení’. Této podmínce se někdy říká *princip superpozice*.

Zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ dané jako

$$\mathbf{f}(\mathbf{x}) = \mathbf{Ax}, \tag{2.5}$$

je lineární, neboť

- $\mathbf{f}(\mathbf{x}+\mathbf{y}) = \mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{y})$
- $\mathbf{f}(\alpha\mathbf{x}) = \mathbf{A}(\alpha\mathbf{x}) = \alpha(\mathbf{Ax}) = \alpha\mathbf{f}(\mathbf{x})$

Prostřední rovnosti plynou z vlastností maticového násobení. Obráceně lze dokázat (udělejte!), že každé lineární zobrazení $\mathbb{R}^n \to \mathbb{R}^m$ lze napsat jako (2.5) pro nějakou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$. Říkáme, že matice $\mathbf{A}$ *reprezentuje* lineární zobrazení.

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

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

tedy $\mathbf{BA}$ je matice složeného zobrazení $\mathbf{g}\circ\mathbf{f}$.

2.4.1 Obraz a nulový prostor

S lineárním zobrazením jsou spjaty dva lineární podprostory, obraz a nulový prostor (jádro). Pokud jde o zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, hovoříme o obrazu a nulovém prostoru matice $\mathbf{A} \in \mathbb{R}^{m\times n}$.

Obraz matice je množina

$$\operatorname{rng} \mathbf{A} = \{\, \mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \,\} \tag{2.6}$$

všech lineárních kombinací sloupců matice. Tedy je to lineární obal sloupců matice. Je to množina $\mathbf{f}(\mathbb{R}^n)$ všech hodnot, jichž může zobrazení $\mathbf{f}$ nabýt. Je to lineární podprostor $\mathbb{R}^m$ s dimenzí $\dim \operatorname{rng} \mathbf{A} = \operatorname{rank} \mathbf{A}$.

Nulový prostor matice je množina

$$\operatorname{null} \mathbf{A} = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0} \,\} \tag{2.7}$$

všech vektorů, které se zobrazí na nulový vektor. Někdy se též nazývá *jádro* (kernel) zobrazení. Nulový prostor je lineární podprostor $\mathbb{R}^n$. Tento podprostor je triviální (obsahuje pouze vektor $\mathbf{0}$) právě tehdy, když matice $\mathbf{A}$ má lineárně nezávislé sloupce. Tedy široká matice má vždy netriviální nulový prostor.

Dimenze obrazu a nulového prostoru jsou svázány vztahem

$$\dim \operatorname{rng} \mathbf{A} + \dim \operatorname{null} \mathbf{A} = n. \tag{2.8}$$

Jak najdeme nulový prostor? Hledáme matici, jejíž sloupce tvoří bázi nulového prostoru matice $\mathbf{A}$. Je to zjevně každá matice $\mathbf{B} \in \mathbb{R}^{n\times k}$ splňující $\mathbf{A}\mathbf{B} = \mathbf{0}$ a $k = \dim \operatorname{null} \mathbf{A} = n - \operatorname{rank} \mathbf{A}$. V Matlabu získáme takovou matici např. příkazem `B=null(A)`.

Příklad 2.3. Dokážeme, že pro každou matici $\mathbf{A}$ platí

$$\operatorname{rank}(\mathbf{A}^T\mathbf{A}) = \operatorname{rank} \mathbf{A}. \tag{2.9}$$

Nejrpve dokážeme, že $\operatorname{null} \mathbf{A} = \operatorname{null}(\mathbf{A}^T\mathbf{A})$:

- Důkaz $\operatorname{null} \mathbf{A} \subseteq \operatorname{null}(\mathbf{A}^T\mathbf{A})$: $\mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0}$.
- Důkaz $\operatorname{null}(\mathbf{A}^T\mathbf{A}) \subseteq \operatorname{null} \mathbf{A}$: $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0} \;\Rightarrow\; \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2 = \mathbf{0} \;\Rightarrow\; \mathbf{A}\mathbf{x} = \mathbf{0}$.

Jelikož jsou nulové prostory matic stejné a n je stejné, podle (2.8) jsou stejné jejich hodnosti.□

2.5 Afinní podprostor a zobrazení

Afinní kombinace vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in V$ je lineární kombinace

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k,$$

ve které $\alpha_1 + \cdots + \alpha_k = 1$.

Afinní obal vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich afinních kombinací. Obecněji, afinní obal (konečné či nekonečné) množiny $X \subseteq V$ je množina

$$\operatorname{aff} X = \{\, \alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k \mid \mathbf{x}_1, \dots, \mathbf{x}_k \in X,\ \alpha_1, \dots, \alpha_k \in \mathbb{R},\ \alpha_1 + \cdots + \alpha_k = 1,\ k \in \mathbb{N} \,\}$$

afinních kombinací všech konečných podmnožin X.

Příklad 2.4. Mějme dva lineárně nezávislé vektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Jejich lineární obal je množina $\operatorname{span}\{\mathbf{x}, \mathbf{y}\} = \{\, \alpha\mathbf{x} + \beta\mathbf{y} \mid \alpha, \beta \in \mathbb{R} \,\}$, tedy rovina procházející těmito dvěma body a počátkem $\mathbf{0}$, tedy celý $\mathbb{R}^2$. Jejich afinní obal je množina

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

což je přímka procházející body $\mathbf{x}, \mathbf{y}$. Nakreslete si vektory $\alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ pro různé hodnoty α!

Podobně, lineární obal dvou lineárně nezávislých vektorů z $\mathbb{R}^3$ je rovina procházející těmito dvěma body a počátkem $\mathbf{0}$, 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. $\square$

Množinu $A \subseteq V$ nazveme **afinní podprostor** lineárního prostoru V, pokud je uzavřená vůči afinním kombinacím (tedy každá afinní kombinace vektorů z A leží v A).

Pokud A je afinní podprostor V a $\mathbf{x}_0 \in A$, pak množina

$$A - \mathbf{x}_0 = \{\, \mathbf{x} - \mathbf{x}_0 \mid \mathbf{x} \in A \,\}$$

je lineární podprostor V. Toto tvrzení dokážeme. 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 $\alpha_1(\mathbf{x}_1 - \mathbf{x}_0) + \cdots + \alpha_k(\mathbf{x}_k - \mathbf{x}_0) \in A - \mathbf{x}_0$, tedy

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

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

Podobně lze dokázat (proveďte), že je-li U lineární podprostor V a $\mathbf{x}_0 \in V$, potom $U + \mathbf{x}_0$ je afinní podprostor V. Tedy afinní podprostor není nic jiného než ‘posunutý’ lineární podprostor. **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**.

Každý afinní podprostor lze také vyjádřit jako množinu řešení nehomogenní soustavy rovnic,

$$\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \,\}.$$

Čtenář si možná všiml, že jsme definovali afinní *podprostor* lineárního prostoru ale už ne afinní *prostor* sám o sobě (bez odkazu k nějakému lineárnímu prostoru). Definice afinního prostoru pomocí axiomů existuje, ale nebudeme ji potřebovat a proto ani uvádět.

Zobrazení $\mathbf{f}\colon U \to V$ z lineárního prostoru U do lineárního prostoru V nazveme **afinní**, pokud (2.4) platí pro všechna $\alpha_1 + \cdots + \alpha_k = 1$, tedy zobrazení afinní kombinace je rovno afinní kombinaci zobrazení. Lze ukázat, že zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ 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

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

Na závěr terminologická poznámka. V lineární algebře znamená ‘lineární zobrazení’ něco jiného než ve zbytku matematiky. 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.

2.6 Cvičení

2.1. Rozhodněte, zda následující množiny tvoří lineární prostor (vždy buď dokažte nebo najděte protipříklad):

a) $\{ \mathbf{x} \mid \mathbf{x}^T\mathbf{y} = 0 \}$ pro daný vektor $\mathbf{y}$
b) Množina všech symetrických matic velikosti $n \times n$.
c) Množina všech singulárních matic velikosti $n \times n$.
d) Množina všech levých inverzí dané úzké matice.

2.2. Které z těchto výroků jsou pravdivé?

a) Pokud $\mathbf{AB}$ má plnou hodnost, pak $\mathbf{A}$ a $\mathbf{B}$ mají plnou hodnost.
b) Pokud $\mathbf{A}$ a $\mathbf{B}$ mají plnou hodnost, pak $\mathbf{AB}$ má plnou hodnost.
c) Pokud $\mathbf{A}$ a $\mathbf{B}$ mají triviální nulový prostor, pak $\mathbf{AB}$ má triviální nulový prostor.
d) Pokud $\mathbf{A}$ a $\mathbf{B}$ jsou úzké s plnou hodností a platí $\mathbf{A}^T\mathbf{B} = \mathbf{0}$, pak matice $[\mathbf{A}\ \mathbf{B}]$ je úzká s plnou hodností.
e) Pokud matice $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$ má plnou hodnost, pak $\mathbf{A}$ i $\mathbf{B}$ mají plnou hodnost.

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.

2.3. Vyřešte tyto rovnice pro neznámou matici $\mathbf{X}$ (předpokládejte, že případná inverze existuje):

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

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

2.5. 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 hodností. 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(n,1)`.

2.6. 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 výhodu oproti počítání $\mathbf{u}$ přímo ze soustavy $\mathbf{Pu} = \mathbf{q}$?

2.7. Které z těchto soustav rovnic jsou lineární? Malá písmena značí vektory, velká matice. Předpokládejte co nejobecnější rozměry matic a vektorů. Jaký je počet rovnic a neznámých v každé soustavě?

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

2.8. Zobrazení $\operatorname{vec}\colon \mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$ ('vektorizace' matice, v Matlabu označeno `A(:)`) je definováno tak, že $\operatorname{vec}\mathbf{A}$ je matice $\mathbf{A}$ přerovnaná po sloupcích do vektoru. *Kroneckerův součin matic* (v Matlabu `kron(A,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

$$\operatorname{vec}(\mathbf{A}\mathbf{B}\mathbf{C}) = (\mathbf{C}^T \otimes \mathbf{A}) \operatorname{vec}\mathbf{B}. \tag{2.10}$$

Použijte tohoto vzorce pro nalezení explicitního řešení následujících soustav rovnic. Neznámá matice je $\mathbf{X}$. Předpokládejte, že počet rovnic je rovný počtu neznámých. Předpokládejte, že matice a vektory mají nejobecnější možné rozměry, aby to dávalo smysl.

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

2.9. Součet prvků na diagonále čtvercové matice se nazývá její *stopa*.

a) Dokažte, že matice $\mathbf{A}\mathbf{B}$ a $\mathbf{B}\mathbf{A}$ mají stejnou stopu.
b) Dokažte, že rovnice $\mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} = \mathbf{I}$ nemá řešení pro žádné $\mathbf{A}, \mathbf{B}$.

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

2.11. Dokažte *Sherman-Morrisonův vzorec* pro inverzi matice $\mathbf{B} = \mathbf{A} - \mathbf{u}\mathbf{v}^T$ ($\mathbf{A}$ je čtvercová regulární):

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

2.12. ($\star$) Uvažujte zobrazení $\mathbf{F}(\mathbf{A}) = (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1}$. Dokažte, že:

a) Pro antisymetrickou matici $\mathbf{A}$ je matice $\mathbf{F}(\mathbf{A})$ ortogonální.
b) Pro ortogonální matici $\mathbf{A}$ je matice $\mathbf{F}(\mathbf{A})$ antisymetrická. Přitom předpokládejte, že $\mathbf{A}$ nemá vlastní číslo -1, jinak by totiž matice $\mathbf{I} + \mathbf{A}$ byla singulární.
c) Zobrazení $\mathbf{F}$ je inverzí sama sebe, tedy $\mathbf{F}(\mathbf{F}(\mathbf{A})) = \mathbf{A}$ pro každé $\mathbf{A}$. Toto má platit pro každou matici $\mathbf{A}$, nejen ortogonální či antisymetrickou.

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

2.13. Výraz $\mathbf{y} \times \mathbf{x}$ označuje vektorový součin vektorů $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. Najděte matici $\mathbf{A}$ tak, aby $\mathbf{y} \times \mathbf{x} = \mathbf{A}\mathbf{x}$. Čemu je rovno $\mathbf{A}^T$? Jakou hodnost má $\mathbf{A}$?

2.14. ($\star$) Matici nazveme surjektivní, pokud lineární zobrazení reprezentované matici je surjektivní, viz §1.2. Které z těchto výroků jsou pravdivé? Dokažte nebo najděte protipříklad.

a) Pokud matice $\mathbf{A}$ a $\mathbf{B}$ jsou surjektivní, pak matice $\mathbf{AB}$ je surjektivní.
b) Pokud matice $\mathbf{A}^2$ je surjektivní, pak $\mathbf{A}$ je surjektivní.
c) Pokud matice $\mathbf{A}^T\mathbf{A}$ je surjektivní, pak $\mathbf{A}$ je surjektivní.
d) Pokud matice $[\mathbf{A}\ \mathbf{0}]$ je surjektivní, pak $\mathbf{A}$ má plnou hodnost.

2.15. Analogicky k pojmu lineární nezávislosti, jak byste definovali 'afinní nezávislost'? Zkuste najít množinu vektorů, která je lineárně závislá a afinně nezávislá. Zkuste to i naopak. V kterém případě je to možné?

Kapitola 3

Skalární součin a ortogonalita

3.1 Skalární součin

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

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$$

(někdy se též značí $\langle \mathbf{x}, \mathbf{y} \rangle$ nebo i jinak). Skalární součin splňuje **Cauchyovu-Schwarzovu nerovnost** $(\mathbf{x} \cdot \mathbf{y})^2 \leq (\mathbf{x} \cdot \mathbf{x})(\mathbf{y} \cdot \mathbf{y})$.

Standardní skalární součin na $\mathbb{R}^n$ indukuje **eukleidovskou normu**

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{{x_1}^2 + \cdots + {x_n}^2}.$$

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

$$\cos \varphi = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}.$$

Eukleidovská norma indukuje **eukleidovskou metriku**

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

Metrika 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 metrickou strukturu.

3.2 Ortogonální vektory a podprostory

Dva vektory nazveme **ortogonální** (kolmé), pokud $\mathbf{u} \cdot \mathbf{v} = 0$, což se značí také $\mathbf{x} \perp \mathbf{y}$. Vektor nazveme **normalizovaný**, pokud má jednotkovou velikost, $\|\mathbf{u}\| = 1$ (tedy $\mathbf{u} \cdot \mathbf{u} = 1$). Množinu vektorů $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ nazveme **ortonormální**, pokud každý vektor je normalizovaný a každá dvojice vektorů je ortogonální.

Ortonormální množina vektorů je lineárně nezávislá. To lze ukázat skalárním vynásobením rovnice $\alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k = \mathbf{0}$ vektorem $\mathbf{u}_i$ (důkaz dokončete!).

Podprostory U a U' lineárního prostoru V se nazývají

- **ortogonální**, je-li $\mathbf{x} \perp \mathbf{x}'$ pro každé $\mathbf{x} \in U$ a $\mathbf{x}' \in U'$. Značíme $U \perp U'$.
- **komplementární**, platí-li zároveň $U \perp U'$ a $\operatorname{span}(U \cup U') = V$. Každý podprostor U má právě jeden komplementární podprostor, který značíme $U^\perp$.

Příklad 3.1. Nechť $V = \mathbb{R}^3$. Dvě na sebe kolmé přímky procházející počátkem jsou ortogonální podprostory. Nejsou komplementární, protože lineární obal jejich sjednocení není celý $\mathbb{R}^3$, ale pouze rovina. Komplementární podprostor k přímce procházející počátkem je rovina procházející počátkem, která je na tuto přímku kolmá. Všimněte si, že tento požadavek určuje rovinu jednoznačně. $\square$

Pro každou matici $\mathbf{A}$ jsou podprostory $\operatorname{null}\mathbf{A}$ a $\operatorname{rng}(\mathbf{A}^T)$ komplementární. To je vidět z toho, že $\operatorname{rng}(\mathbf{A}^T)$ je množina všech lineárních kombinací řádků matice $\mathbf{A}$ a $\operatorname{null}\mathbf{A}$ je množina všech vektorů kolmých na všechny řádky matice $\mathbf{A}$. Důkaz dokončete!

3.3 Ortogonální matice

Předpodkládejme, že vektory tvoří sloupce matice $\mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_k\end{bmatrix} \in \mathbb{R}^{n\times k}$. Ortonormalitu sloupců lze nyní popsat podmínkou

$$\mathbf{U}^T\mathbf{U} = \mathbf{I},$$

tedy $\mathbf{U}^T$ je levá inverze $\mathbf{U}$. Z lineární nezávislosti sloupců plyne, že nezbytně $k \le n$. Lineární zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{U}\mathbf{x}$ *zachovává standardní skalární součin*, neboť

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

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

Pro obdélníkovou matici s ortonormálními sloupci neexistuje standardní jméno. Pokud je matice $\mathbf{U}$ čtvercová ($k = n$), následující podmínky jsou ekvivalentní (proč?):

$$\mathbf{U}^T\mathbf{U} = \mathbf{I} \quad \Longleftrightarrow \quad \mathbf{U}^{-1} = \mathbf{U}^T \quad \Longleftrightarrow \quad \mathbf{U}\mathbf{U}^T = \mathbf{I}.$$

V tom případě se $\mathbf{U}$ nazývá **ortogonální matice** (z historických důvodů se neříká 'ortonormální matice'). Z podmínek pak plyne $(\det\mathbf{U})^2 = 1$, mohou tedy nastat dva případy:

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

Příklad 3.2. 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, že $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$ a $\det\mathbf{U} = 1$.

Permutační matice je čtvercová matice, jejíž sloupce jsou permutované ('přeházené') vektory standardní báze. Např. matice $\begin{bmatrix}\mathbf{e}_3 & \mathbf{e}_1 & \mathbf{e}_2\end{bmatrix}$. Permutační matice je zjevně ortogonální. Její determinant je rovný znaménku permutace. $\square$

Ortonormální množiny vektorů mají mnoho výhod:

- Je to 'nejvíce lieárně nezávislá' množina vektorů.
- Tvoří 'nejhezčí' báze lineárních podprostorů.
- Inverze ortonormální matice se spočítá triviálně (transpozicí) a nevznikají u toho numerické (zaokrouhlovací) chyby. Naproti tomu inverze obecné matice má složitost $O(n^3)$ a mohou vznikat značné zaokrouhlovací chyby.

3.4 Gramm-Schmidtova ortonormalizace

(Tato část je nepovinná.)

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

- $\mathbf{q}_1, \ldots, \mathbf{q}_n$ jsou ortonormální,
- Pro každé $k = 1, \ldots, n$ platí $\operatorname{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_k\} = \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$.

Myšlenka algoritmu je jednoduchá. Předpokládejme, že již máme vektory $\mathbf{q}_1, \ldots, \mathbf{q}_{k-1}$ s popsanými vlastnostmi. K vektoru $\mathbf{a}_k$ přičteme takovou lineární kombinaci vektorů $\mathbf{q}_1, \ldots, \mathbf{q}_{k-1}$, aby se stal na všechny z nich ortogonální. Poté tento vektor normalizujeme. Tedy

$$\mathbf{q}_k := \mathbf{a}_k - \sum_{j=1}^{k-1} r_{jk} \mathbf{q}_j, \quad \mathbf{q}_k := \frac{\mathbf{q}_k}{\|\mathbf{q}_k\|}. \tag{3.1}$$

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

Jak najdeme koeficienty r_{jk}? Ze (3.1) vyplývá, že

$$\mathbf{a}_k = \sum_{j=1}^{k} r_{jk} \mathbf{q}_j. \tag{3.2}$$

Máme zde navíc koeficient r_{kk}, který reprezentuje změnu vektoru $\mathbf{q}_k$ normalizací. Vztah (3.2) nám dovoluje spočítat koeficienty r_{jk} z požadavku na ortonormalitu vektorů $\mathbf{q}_1, \ldots, \mathbf{q}_k$. Jeho vynásobením vektorem $\mathbf{q}_j$ dostaneme $r_{jk} = \mathbf{q}_j \cdot \mathbf{a}_k$.

3.5 QR rozklad

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

$$\mathbf{A} = \mathbf{Q}\mathbf{R}, \tag{3.3}$$

kde $\mathbf{Q}$ má ortonormální sloupce ($\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$) a $\mathbf{R}$ je horní trojúhelníková (tj. $r_{ij} = 0$ pro každé $i > j$). Existují dvě verze:

- V **redukovaném QR rozkladu** je $\mathbf{Q} \in \mathbb{R}^{m \times n}$ a $\mathbf{R} \in \mathbb{R}^{n \times n}$. V Matlabu `[Q,R]=qr(A,0)`.
- V **úplném QR rozkladu** je $\mathbf{Q} \in \mathbb{R}^{m \times m}$ a $\mathbf{R} \in \mathbb{R}^{m \times n}$. V Matlabu `[Q,R]=qr(A)`. Úplný QR rozklad lze napsat jako

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix}, \tag{3.4}$$

kde $\mathbf{A} = \mathbf{Q}_1\mathbf{R}_1$ je redukovaný QR rozklad. Tedy máme navíc matici $\mathbf{Q}_2$ s ortonormálními sloupci, které jsou kolmé na sloupce $\mathbf{Q}_1$. Podprostory $\operatorname{rng}\mathbf{Q}_1$ a $\operatorname{rng}\mathbf{Q}_2$ jsou tedy komplementární.

QR rozklad je základním stavebním kamenem mnoha dalších algoritmů numerické lineární algebry vektorů. Je také užitečný pro řešení lineárních rovnic. Řešme např. soustavu $\mathbf{Ax} = \mathbf{b}$ s regulární čtvercovou maticí $\mathbf{A}$. Rozložíme $\mathbf{A} = \mathbf{QR}$ a vynásobíme soustavu zleva $\mathbf{Q}^T$, což dá $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$. Toto je ekvivalentní úprava, neboť $\mathbf{Q}$ je regulární. Ale protože je $\mathbf{R}$ trojúhelníková, tato soustava se levně vyřeší zpětnou substitucí.

QR rozklad lze počítat např. vylepšenou Gramm-Schmidtovou ortonormalizací. Rovnice (3.2) se dá psát v maticovém tvaru jako (3.3), kde vektory $\mathbf{a}_1, \ldots, \mathbf{a}_n$ jsou sloupce matice $\mathbf{A}$, vektory $\mathbf{q}_1, \ldots, \mathbf{q}_n$ jsou sloupce $\mathbf{Q}$ a $\mathbf{R}$ je horní trojúhelníková s prvky $r_{jk} = \mathbf{q}_i \cdot \mathbf{a}_k$. Redukovaný QR rozklad pak vznikne vylepšeními tohoto algoritmu, které jednak zmenší zaokrouhlovací chyby a jednak dovolí lineární závislost sloupců $\mathbf{A}$.

3.6 Cvičení

3.1. Dokažte, že součin ortogonálních matic je ortogonální matice.

3.2. Pro jaké n je matice $\operatorname{diag}(-\mathbf{1})$ (tedy diagonální matice s mínus jedničkami na diagonále) rotační?

3.3. Počet nezávislých parametrů (stupňů volnosti) ortogonální matice $n \times n$ se získá jako rozdíl počtu prvků matice (n^2) a počtu nezávislých rovnic v podmínce $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. Neformálně řečeno udává, kolika ‘knoflíky’ můžeme nezávisle ‘kroutit’ při rotaci v n-rozměrném prostoru. Jaké je toto číslo pro $n = 2, 3, 4$? Najděte vzorec pro obecné n.

3.4. Pro $\|\mathbf{v}\| = 1$ je $\mathbf{H} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^T$ známa jako *Householderova matice*. Transformace $\mathbf{Hx}$ je zrcadlení vektoru $\mathbf{x}$ kolem nadroviny s normálovým vektorem $\mathbf{v}$, proto se $\mathbf{H}$ také někdy nazývá *elementární reflektor*.

   a) Ukažte, že $\mathbf{H} = \mathbf{H}^T$ a $\mathbf{H}^T\mathbf{H} = \mathbf{I}$ (tj. matice je symetrická a ortogonální).
   b) Z těchto dvou vlastností vyplývá $\mathbf{HH} = \mathbf{I}$. Co to říká o transformaci $\mathbf{Hx}$?
   c) Ukažte, že $\det \mathbf{H} = -1$.
   d) Co je $\mathbf{Hv}$? Co je $\mathbf{Hx}$, kde $\mathbf{v}^T\mathbf{x} = 0$? Ukažte algebraicky a odůvodněte geometricky.

3.5. ($\star$) RQ rozklad rozloží matici $\mathbf{A} = \mathbf{RQ}$, kde $\mathbf{R}$ je horní trojúhelníková a $\mathbf{Q}$ je ortogonální. Jak spočítáte RQ rozklad, máte-li počítačovou implementaci QR rozkladu?

Kapitola 4

Metoda nejmenších čtverců

Řešme nehomogenní soustavu lineárních rovnic

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{4.1}$$

kde $\mathbf{A} \in \mathbb{R}^{m\times n}$. Soustava má řešení právě tehdy, když $\mathbf{b} \in \operatorname{rng} \mathbf{A}$, tedy je-li $\mathbf{b}$ lineární kombinací sloupců $\mathbf{A}$ (Frobeniova věta). V opačném případě je soustava **přeurčená**. To nastane typicky pro $m > n$ (tedy rovnic je více než neznámých) – tato podmínka však není ani nutná ani postačující.

Řešme přeurčenou soustavu přibližně. Hledejme takové $\mathbf{x}$, že vzdálenost mezi body $\mathbf{A}\mathbf{x}$ a $\mathbf{b}$ je co nejmenší, když už nemůže být nulová:

$$\min\{\, \|\mathbf{A}\mathbf{x} - \mathbf{b}\| \mid \mathbf{x} \in \mathbb{R}^n \,\}, \tag{4.2}$$

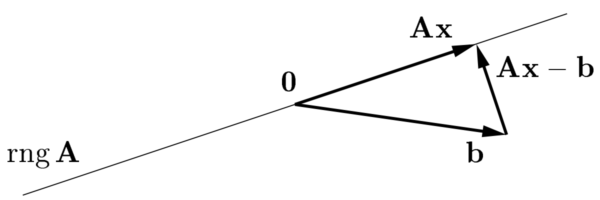
Je jasné, že místo normy $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|$ můžeme minimalizovat její čtverec $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$. Protože to vlastně znamená minimalizaci součtu čtverců reziduí $\mathbf{a}_i^T\mathbf{x} - b_i$ (kde $\mathbf{a}_i^T$ jsou řádky $\mathbf{A}$), mluvíme o přibližném řešení **ve smyslu nejmenších čtverců** (*least squares solution*).

Příklad 4.1. Soustava třech rovnic o dvou neznámých

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

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

Úlohu lze řešit čistě geometrickou úvahou. Pokud vzdálenost $\mathbf{A}\mathbf{x}$ a $\mathbf{b}$ má být minimální, musí být vektor $\mathbf{A}\mathbf{x} - \mathbf{b}$ kolmý na prostor $\operatorname{rng} \mathbf{A}$, tedy na každý sloupec matice $\mathbf{A}$ (viz obrázek).



To lze zapsat jako $\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{0}$, tedy

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}. \tag{4.3}$$

Soustava (4.3) se proto nazývá **normální rovnice**. Je to soustava n rovnic o n neznámých (to je příjemné, když $m \gg n$, což se v praktických úlohách často stává).

Věta 4.1. *Rovnice (4.3) má řešení pro každé* $\mathbf{A}$ *a* $\mathbf{b}$*.*

Důkaz. Tvrdíme, že

$$\operatorname{rank}\mathbf{A} \overset{(a)}{=} \operatorname{rank}\mathbf{A}^T\mathbf{A} \overset{(b)}{\leq} \operatorname{rank}\mathbf{A}^T\begin{bmatrix}\mathbf{A} & \mathbf{b}\end{bmatrix} \overset{(c)}{\leq} \operatorname{rank}\mathbf{A}.$$

Rovnost (a) byla dokázána v Příkladu 2.3. Nerovnost (b) je evidentní, neboť přidáním sloupce nemůže hodnost matice $\mathbf{A}^T\mathbf{A}$ klesnout. Nerovnost (c) plyne z (2.2). Jelikož vpravo i vlevo je stejné číslo, musí být $\operatorname{rank}\mathbf{A}^T\mathbf{A} = \operatorname{rank}\mathbf{A}^T\begin{bmatrix}\mathbf{A} & \mathbf{b}\end{bmatrix}$ a podle Frobeniovy věty má soustava (4.3) řešení. □

Když $\operatorname{rank}\mathbf{A} = n$, matice $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulární a soustavu můžeme řešit pomocí inverze:

$$\mathbf{x} = \mathbf{A}^+\mathbf{b}, \quad \text{kde} \quad \mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T. \tag{4.4}$$

Matice $\mathbf{A}^+$ se nazývá **pseudoinverze** matice $\mathbf{A}$. Je to jedna z levých inverzí matice $\mathbf{A}$, neboť

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

Lze dokázat, že pseudoinverze $\mathbf{A}^+$ má mezi všemi levými inverzemi $\mathbf{B}$ matice $\mathbf{A}$ nejmenší Frobeniovu normu

$$\|\mathbf{B}\|_\mathrm{F} = \Big(\sum_{i,j} {b_{ij}}^2\Big)^{1/2}.$$

Pokud $\operatorname{rank}\mathbf{A} < n$, je matice $\mathbf{A}^T\mathbf{A}$ singulární a řešení pomocí inverze nemůžeme použít. V tom případě rovnice (4.3), a tedy i úloha (4.2), mají nekonečně mnoho (afinní podprostor) řešení $\mathbf{x}$. To se obvykle nestává – může to znamenat, že jsme úlohu špatně zformulovali.

Normální rovnice lze také řešit pomocí redukovaného QR rozkladu. Nechť $\operatorname{rank}\mathbf{A} = n$ a $\mathbf{A} = \mathbf{QR}$. Po dosazení do (4.3) 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 $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$. To řešíme zpětnou substitucí.

Dosaďme $\mathbf{x}$ spočítané vztahem (4.4) do levé strany (4.1):

$$\mathbf{Ax} = \mathbf{Pb}, \quad \text{kde} \quad \mathbf{P} = \mathbf{AA}^+ = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T. \tag{4.5}$$

Matice $\mathbf{P} \in \mathbb{R}^{m\times m}$ má význam *ortogonální projekce* na podprostor $\operatorname{rng}\mathbf{A}$. *Hodnota* minima problému (4.2) je $\|\mathbf{Pb} - \mathbf{b}\| = \|(\mathbf{P} - \mathbf{I})\mathbf{b}\|$.

4.1 Statistické odůvodnění metody

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

$$\mathbf{y} = \mathbf{Ax} + \boldsymbol{\varepsilon},$$

kde $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m)$ jsou náhodné proměnné modelující chyby měření $\mathbf{y}$. Metoda nejmenších čtverců činí dva předpoklady:

- Náhodné proměnné ε_i mají *normální rozdělení* s nulovou střední hodnotou a směrodatnou odchylkou σ,

$$p(\varepsilon_i) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\varepsilon_i^2/(2\sigma^2)}.$$

- Náhodné proměnné $\varepsilon_1, \ldots, \varepsilon_m$ jsou na sobě nezávislé. Tedy sdružená pravděpodobnost je rovna součinu

$$p(\varepsilon_1, \ldots, \varepsilon_m) = \prod_{i=1}^{m} p(\varepsilon_i) = \prod_{i=1}^{m} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-{\varepsilon_i}^2/(2\sigma^2)}. \tag{4.6}$$

Dále aplikujeme *princip maxima věrohodnosti*. Ten říká, že parametry $\mathbf{x}$ se mají najít tak, aby pravděpodobnost (4.6) byla maximální. Je pohodlnější minimalizovat její záporný logaritmus

$$-\log p(\varepsilon_1, \ldots, \varepsilon_m) = -\sum_{i=1}^{m} \log p(\varepsilon_i) = \sum_{i=1}^{m} \Big[\frac{{\varepsilon_i}^2}{2\sigma^2} + \log(\sigma\sqrt{2\pi})\Big].$$

Jelikož σ je konstanta, je to totéž jako minimalizovat $\sum_i {\varepsilon_i}^2 = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2$.

4.2 Použití na regresi

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

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

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

$$\min_{\mathbf{x}\in\mathbb{R}^n} \sum_{i=1}^{m} [f(t_i, \mathbf{x}) - y_i]^2. \tag{4.7}$$

Zvolme regresní funkci jako lineární kombinaci

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

daných *bázových funkcí* $\varphi_i \colon T \to \mathbb{R}$. Pak se úloha (4.7) dá psát jako $\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2$, kde prvky matice $\mathbf{A}$ jsou $a_{ij} = \varphi_j(t_i)$.

Příklad 4.2. *Polynomiální regrese.* Nechť $X = \mathbb{R}$ a $\varphi_i(t) = t^{i-1}$. Pak regresní funkce je polynom stupně $n-1$,

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

Speciálně pro $n = 1$ úloha (4.7) zní $\min_x \sum_i (y_i - x)^2$. Řešením je známý aritmetický průměr $x = \frac{1}{m}\sum_{i=1}^{m} y_i$. □

4.3 Cvičení

4.1. V Matlabu je řešení soustavy $\mathbf{Ax} = \mathbf{b}$, ať už přesně určené nebo přeurčené, implementováno v operátorech \ (zpětné lomítko) a / (lomítko). Pochopte všechny funkce těchto operátorů pomocí studia příkazů `help mldivide` a `help mrdivide`.

4.2. **Projekcí** se v lineární algebře rozumí každé lineární zobrazení $\mathbf{f}(\mathbf{y}) = \mathbf{Py}$, které splňuje $\mathbf{f}(\mathbf{f}(\mathbf{y})) = \mathbf{f}(\mathbf{y})$, tedy $\mathbf{P}^2 = \mathbf{P}$. To vyjadřuje pochopitelný požadavek, že když jednou vektor promítneme, tak další promítnutí ho již nezmění. Projekce obecně nemusí být ortogonální. Projekce je ortogonální, když navíc platí $\mathbf{P}^T = \mathbf{P}$. Zkontrolujte, že matice (4.5) vyhovuje uvedeným rovnostem.

4.3. V prknu je n děr o souřadnicích $x_1, \ldots, x_n \in \mathbb{R}$, všechny v jedné přímce. Naměříme metrem vzdálenosti $d_{ij} = x_j - x_i$ pro vybrané dvojice $(i, j) \in E$, kde množina $E \subseteq \{1, \ldots, n\}^2$ je dána. Přitom dvojice jsou vybrané tak, že vždy $x_j > x_i$.

   a) Zformulujte odhad $x_1, \ldots, x_n$ jako úlohu (4.2), tj. najděte matici $\mathbf{A}$ a vektor $\mathbf{b}$.
   b) Jaká je nejmenší možná velikost množiny E, aby úloha měla smysl?
   c) Lze dosáhnout, aby měla matice $\mathbf{A}$ plnou hodnost? Pokud ne, jak byste úlohu změnili, aby měla?

4.4. Zformulujte nalezení příčky mimoběžek v $\mathbb{R}^3$ jako úlohu ve tvaru (4.2) (udejte přesný tvar matice $\mathbf{A}$ a vektoru $\mathbf{b}$). Úlohu vyřešte a dokažte, že příčka je na mimoběžky kolmá.

4.5. Máme množinu m přímek (afinních podprostorů dimenze 1) v prostoru $\mathbb{R}^n$, kde i-tá přímka je množina $\{\, \mathbf{a}_i + t\mathbf{s}_i \mid t \in \mathbb{R} \,\}$. Najděte bod, jehož součet čtverců vzdáleností k přímkám je minimální. Jak byste to zobecnili na případ, kdy místo m přímek máme m afinních podprostorů s dimenzemi $d_1, \ldots, d_m$?

4.6. 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 $b_i \in \mathbb{R}$. Najděte bod, který minimalizuje součet čtverců vzdáleností od jednotlivých přímek. Nápověda: Jak se počítá vzdálenost bodu od nadroviny, tedy např. od přímky v rovině?

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

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

kde w_i jsou nezáporné váhy. Napište funkci v maticovém tvaru (nápověda: sdružte čísla w_i do diagonální matice). Za jakých podmínek má úloha řešení? Jak se řešení spočítá?

Kapitola 5

Vlastní čísla a kvadratické formy

5.1 Vlastní čísla a vektory

Definice 5.1. *Nechť pro čtvercovou matici* $\mathbf{A} \in \mathbb{R}^{n \times n}$, *vektor* $\mathbf{v} \neq \mathbf{0}$ *a skalár* $\lambda \in \mathbb{C}$ *platí*

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}. \tag{5.1}$$

Pak λ *se nazývá* **vlastní číslo** *matice a* $\mathbf{v}$ **vlastní vektor** *matice příslušný vlastnímu číslu* λ.

Rovnici (5.1) lze přepsat jako

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}. \tag{5.2}$$

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 řešením rovnice

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{5.3}$$

Levá strana této rovnice se nazývá **charakteristický polynom**. Vlastní vektory příslušné vlastnímu číslo λ pak spočítáme ze soustavy (5.2).

Příklad 5.1. Najděte vlastní čísla matice $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Charakteristická rovnice zní

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

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

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

□

Charakteristický polynom má stupeň n (dokažte z definice determinantu!). Má tedy n kořenů, z nichž některé mohou násobné (tj. stejné). V tomto smyslu má každá matice n vlastních čísel, z nichž některá mohou být násobná. Vlastní čísla mohou být reálná nebo komplexní. Množině vlastních čísel matice (či lineárního zobrazení) se někdy říká její **spektrum**.

Všechny vlastní vektory příslušné danému vlastnímu číslu tvoří lineární podprostor $\mathbb{R}^n$. Speciálně, velikost vlastních vektorů nehraje roli a je proto zvykem je normalizovat, $\|\mathbf{v}\| = 1$.

Celkové množství všech lineárně nezávislých vlastních vektorů může být nejvýše n (protože více než n vektorů v $\mathbb{R}^n$ je vždy lineárně závislých), může jich být ale i méně. Vlastní vektory příslušné různým vlastním číslům jsou lineárně nezávislé (dokažte!).

Rovnice (5.1) lze napsat pro všechna vlastní čísla najednou jako

$$\mathbf{A}\begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix},$$

neboli

$$\mathbf{AV} = \mathbf{VD}. \tag{5.4}$$

Diagonální matice $\mathbf{D}$ má na diagonále vlastní čísla a sloupce čtvercové matice $\mathbf{V}$ jsou vlastní vektory. Přitom se předpokládá, že sloupce $\mathbf{V}$ jsou ze všech vlastních vektorů vybrány tak, aby hodnost $\mathbf{V}$ byla největší možná. Pokud je $\mathbf{V}$ regulární, (5.4) lze psát jako

$$\mathbf{A} = \mathbf{VDV}^{-1}. \tag{5.5}$$

Vztahu (5.5) se pak říká **rozklad matice podle vlastních čísel** nebo **spektrální rozklad**.

Věta 5.1. *Nechť matice $\mathbf{A}$ rozměru $n \times n$ je symetrická. Pak všechna její vlastní čísla jsou reálná a lze najít n navzájem ortogonálních vlastních vektorů.*

Této větě se někdy říká **spektrální věta**. Podle ní pro každou symetrickou $\mathbf{A}$ je ve vztahu (5.4) matice $\mathbf{D}$ reálná a $\mathbf{V}$ může být zvolena jako ortogonální, $\mathbf{V}^{-1} = \mathbf{V}^T$. Tedy

$$\mathbf{A} = \mathbf{VDV}^T. \tag{5.6}$$

Jak se počítají vlastní čísla a vektory? Charakteristický polynom je hlavně teoertický nástroj a přímé hledání jeho kořenů není vhodné pro numerický výpočet. Pro větší matice se vlastní čísla a vektory počítají iteračními algoritmy. Navíc, pro matice různého typu jsou vhodné různé algoritmy – např. symetrické matice je výpočet významně snazší. Matlabská funkce `[V,D]=eig(A)` spočítá matice $\mathbf{V}$ a $\mathbf{D}$ splňující (5.4).

5.2 Kvadratická forma

Kvadratická forma je funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ daná vzorcem

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

kde $\mathbf{A}$ je matice velikosti $n \times n$.

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

$$\mathbf{A} = \underbrace{\frac{\mathbf{A}+\mathbf{A}^T}{2}}_{\mathbf{A}_+} + \underbrace{\frac{\mathbf{A}-\mathbf{A}^T}{2}}_{\mathbf{A}_-}.$$

Ale $\mathbf{x}^T\mathbf{A}\mathbf{x} = (\mathbf{x}^T\mathbf{A}\mathbf{x})^T = \mathbf{x}^T\mathbf{A}^T\mathbf{x}$, neboť transpozice skaláru je tentýž skalár. Z toho plyne

$$\mathbf{x}^T\mathbf{A}_-\mathbf{x} = 0, \quad \mathbf{x}^T\mathbf{A}_+\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x}.$$

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

Definice 5.2. *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 dobrý smysl pro libovoné čtvercové matice, je zvykem hovořit o těchto vlastnostech jen pro symetrické matice. Někdy se vlastnosti definují ne pro matici, ale abstraktněji pro kvadratickou formu.

Má-li kvadratická forma na množině $\mathbb{R}^n$ extrém, bude se nabývat v počátku $\mathbf{x} = \mathbf{0}$ (i když možná i jinde). Definice 5.2 ukazuje, 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.

Věta 5.2. *Symetrická matice je*

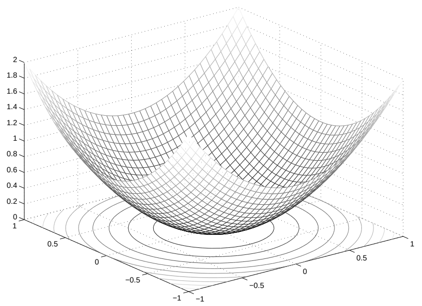
- *pozitivně [negativně] semidefinitní, právě když má všechna vlastní čísla nezáporná [nekladná]*
- *pozitivně [negativně] definitní, právě když má všechna vlastní čísla kladná [záporná]*
- *indefinitní, právě když má alespoň jedno kladné a alespoň jedno záporné vlastní číslo.*

Důkaz. Z rozkladu podle vlastních čísel (5.6) máme

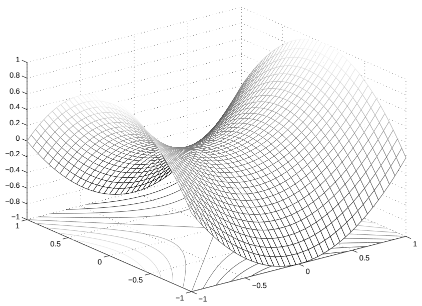
$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{x} = \mathbf{y}^T\mathbf{D}\mathbf{y} = \lambda_1 {y_1}^2 + \cdots + \lambda_n {y_n}^2. \tag{5.7}$$

kde $\mathbf{x} = \mathbf{V}\mathbf{y}$. Substituce $\mathbf{x} = \mathbf{V}\mathbf{y}$ tedy diagonalizovala matici kvadratické formy. Protože $\mathbf{V}$ je regulární, definitnost matice $\mathbf{A}$ je stejná jako definitnost matice $\mathbf{D}$. Ale protože $\mathbf{D}$ je diagonální, její definitnost je okamžitě patrná ze znamének čísel λ_i. Např. výraz (5.7) je nezáporný pro každé $\mathbf{y}$ právě tehdy, když všechna λ_i jsou nezáporná. □

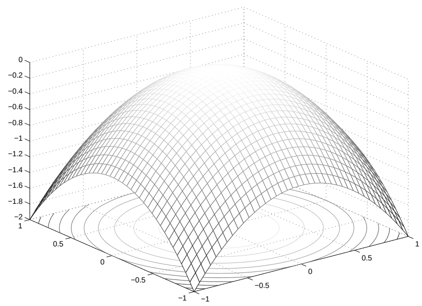
Když jsou všechna λ_i kladná ($\mathbf{A}$ pozitivně definitní), funkce $g(\mathbf{y}) = \mathbf{y}^T\mathbf{D}\mathbf{y}$ vypadá jako 'údolí'. Když jsou všechna λ_i záporná ($\mathbf{A}$ negativně definitní), funkce vypadá jako 'kopec'. Když jsou některá kladná a některá záporná ($\mathbf{A}$ indefinitní), tvarem je 'sedlo':



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



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

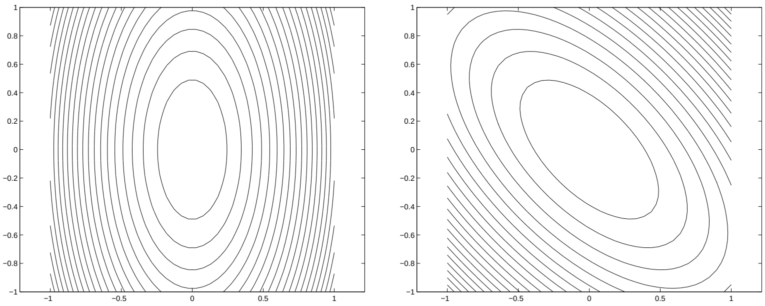


$g(y_1, y_2) = -y_1^2 - y_2^2$

Protože však $\mathbf{V}$ je ortogonální, transformace $\mathbf{x} = \mathbf{V}\mathbf{y}$ je pouhá isometrie. Např. nechť

$$\mathbf{D} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{V} = \frac{1}{2}\begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & \sqrt{2} \end{bmatrix},$$

tedy $\mathbf{V}$ reprezentuje otočení o $\pi/4$. Obrázek ukazuje vrstevnice původní a otočené funkce:



$g(\mathbf{y}) = \mathbf{y}^T\mathbf{D}\mathbf{y}$ $\qquad$ $f(\mathbf{x}) = g(\mathbf{V}^T\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$

5.3 Kvadratická funkce

Obecná **kvadratická funkce** více proměnných má tvar

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c, \tag{5.8}$$

oproti kvadratické formě tedy přibyl lineární a konstantní člen. Opět předpokládáme $\mathbf{A}^T = \mathbf{A}$.

Někdy lze tuto funkci přepsat jako

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = (\mathbf{x} - \mathbf{x}_0)^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0) + y_0. \tag{5.9}$$

Výraz na pravé straně je kvadratická forma s počátkem $\mathbf{x}_0$, plus konstanta. Této úpravě se říká ‘doplnění na čtverec’. Znáte ji pro případ, kdy $\mathbf{A}, \mathbf{b}, \mathbf{x}$ jsou skaláry – tak se na základní škole odvozuje známý vzorec pro kořeny kvadratické rovnice jedné proměnné. Spočtěme $(\mathbf{x}_0, y_0)$ z $(\mathbf{A}, \mathbf{b}, c)$. Roznásobením pravé strany dostaneme

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

kde jsme použili $\mathbf{x}^T\mathbf{A}\mathbf{x}_0 + \mathbf{x}_0^T\mathbf{A}\mathbf{x} = 2\mathbf{x}^T\mathbf{A}\mathbf{x}_0$. Porovnáním s levou stranou máme

$$\mathbf{b} = -2\mathbf{A}\mathbf{x}_0, \quad c = \mathbf{x}_0^T\mathbf{A}\mathbf{x}_0 + y_0,$$

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

Pokud je doplnění na čtverec možné, vyšetření extrémů kvadratické funkce se neliší od vyšetření extrémů kvadratické formy – rozdíl je jen v posunutí $\mathbf{x}_0$. Pokud doplnění možné není, kvadratická funkce extrém nemá ani když $\mathbf{A}$ je pozitivně či negativně semidefinitní. Promyslete, jak funkce v tom případě vypadá!

5.4 Cvičení

5.1. Spočítejte vlastní čísla a vlastní vektory matic $\begin{bmatrix} 1 & 2 \\ -4 & -2 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}$.

5.2. Napište rovnici, jejímiž kořeny jsou vlastní čísla matice $\begin{bmatrix} 2 & 0 & 3 \\ 0 & -2 & -1 \\ 3 & -1 & 2 \end{bmatrix}$.

5.3. Najděte všechna vlastní čísla a vlastní vektory (a) nulové, (b) jednotkové, (c) diagonální matice. Najděte vlastní čísla trojúhelníkové matice.

5.4. Ukažte, že $\lambda_1 + \cdots + \lambda_n = \operatorname{trace} \mathbf{A}$ a $\lambda_1 \times \cdots \times \lambda_n = \det \mathbf{A}$.

5.5. Známe vlastní čísla a vektory matice $\mathbf{A}$. Jaké jsou vlastní čísla a vektory matice $\mathbf{A} + \alpha\mathbf{I}$?

5.6. Řekli jsme, že hledání kořenů charakteristického polynomu (5.3) není vhodný způsob na hledání vlastních čísel. Naopak: hledání kořenů libovolného polynomu jde převést na hledání vlastních čísel matice, která se nazývá *doprovodná matice* polynomu. Odvoďte tvar této matice. Ověřte v Matlabu pro různé polynomy.

5.7. 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.

5.8. Ve Cvičení 4.2 jsme definovali projekci jako matici $\mathbf{P}$ splňující $\mathbf{P}^2 = \mathbf{P}$. Geometrickou úvahou najděte aspoň jedno vlastní číslo a příslušný vlastní vektor projekce.

5.9. Geometrickou úvahou najděte aspoň dva vlastní vektory a příslušná vlastní čísla Householderovy matice ze Cvičení 3.4.

5.10. Čemu je rovno $\mathbf{A}^n$, kde $\mathbf{A}$ je symetrická matice?

5.11. Ukažte, že vlastní čísla antisymetrické matice jsou nula nebo čistě imaginární.

5.12. ($\star$) Ukažte, že dvě čtvercové matice komutují, právě když mají stejné vlastní vektory.

5.13. ($\star$) Nechť $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{n\times m}$. Ukažte, že nenulová vlastní čísla matic $\mathbf{AB}$ a $\mathbf{BA}$ jsou stejná.

5.14. Pro dané matice určete, zda jsou pozitivně/negativně (semi)definitní nebo indefinitní:

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

5.15. Najdi minimum kvadratické funkce

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

b) $f(\mathbf{x}) = \mathbf{x}^T \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2 & -1 \end{bmatrix} \mathbf{x}$ Nápověda: Doplňte na úplný čtverec.

5.16. 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.

Nápověda: Je matice symetrická?

5.17. 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).

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

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

5.20. Musí mít pozitivně semidefinitní matice na diagonále nezáporné prvky? V případě kladné odpovědi dokažte, v případě záporné odpovědi najděte protipříklad.

5.21. 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 $\mathbf{A}\mathbf{B} \succeq 0$.
e) Pokud $\mathbf{A} \succ 0$, potom $\mathbf{A}^{-1} \succ 0$.
f) ($\star$) Pokud $\mathbf{A} \succeq 0$ a $\mathbf{B} \succeq 0$, potom $\mathbf{A}\mathbf{B}\mathbf{A} \succeq 0$.

Které z těchto tvrzení doopravdy platí? Dokažte nebo najděte protipříklady.

5.22. Uvažujme náhodnou čtvercovou matici, jejíž prvky jsou nezávislá náhodná čísla z normálního rozdělení s nulovou střední hodnotou a jednotkovou variancí. Takovou matici získáme v Matlabu příkazem `A=randn(n)`. Budeme-li takto generovat velké množství matic, kolik mezi nimi bude pozitivně definitních, kolik pozitivně semidefinitních, a kolik indefinitních? Odůvodněte. Zkuste v Matlabu pro konečné vzorky matic.

Kapitola 6

Množiny a zobrazení v eukleidovských prostorech

6.1 Minimum a infimum

Množina $\mathbb{R}$ reálných čísel je přirozeně obdařena *úplným uspořádáním*, které značíme $\leq$. Pro množinu $Y \subseteq \mathbb{R}$ definujme:

- **Dolní závora** (také **dolní mez**, *lower bound*) množiny Y je každý prvek $a \in \mathbb{R}$, pro který je $a \leq y$ pro všechna $y \in Y$.
- **Nejmenší prvek** (krátce **minimum**) množiny Y je její dolní závora, která leží v množině. Pokud existuje, je určena jednoznačně. Značíme $a = \min Y$.
- **Infimum** množiny Y je největší dolní závora množiny Y. Značíme $a = \inf Y$.

Horní závora, největší prvek (maximum, $\max Y$) a supremum ($\sup Y$) se definují analogicky.

Minimum či maximum podmnožiny reálných čísel nemusí existovat. Je hlubokou vlastností reálných čísel, že v nich existuje infimum [supremum] každé zdola [shora] omezené podmnožiny. Tato vlastnost se nazývá *úplnost*. Pokud Y je zdola [shora] neomezená, definujeme $\inf Y = -\infty$ [$\sup Y = +\infty$]. Pro prázdnou množinu definujeme $\inf \emptyset = +\infty$ a $\sup \emptyset = -\infty$.

Příklad 6.1.

1. Množina všech horních závor intervalu $\langle 0, 1)$ je $\langle 1, +\infty)$.
2. Množina všech horních závor množiny $\mathbb{R}$ je $\emptyset$.
3. Množina všech horních závor množiny $\emptyset$ je $\mathbb{R}$.
4. Interval $\langle 0, 1)$ nemá největší prvek, ale má supremum 1.
5. Minimum množiny $\{\, 1/n \mid n \in \mathbb{N} \,\}$ neexistuje, ale infimum je rovno 0.
6. Maximum množiny $\{\, x \in \mathbb{Q} \mid x^2 \leq 2 \,\} = \langle -\sqrt{2}, \sqrt{2} \rangle \cap \mathbb{Q}$ neexistuje, ale supremum je rovno $\sqrt{2}$.
7. $\max\{1, 2, 3\} = \sup\{1, 2, 3\} = 3$ (minimum a maximum každé konečné množiny existují a jsou rovny infimu a supremu)
8. $\max \mathbb{R}$ neexistuje. $\square$

6.2 Podmnožiny eukleidovského prostoru

Uvažujme prostor $\mathbb{R}^n$ vybavený eukleidovskou metrikou $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Existence metriky nám dovolí zkoumat topologicko-metrické vlastnosti množin v $\mathbb{R}^n$.

Pro $\varepsilon > 0$ a $\mathbf{x} \in \mathbb{R}^n$ se množina

$$U_\varepsilon(\mathbf{x}) = \{\, \mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| < \varepsilon \,\}$$

nazývá **ε-okolí** bodu $\mathbf{x}$. Je to koule (bez hranice) se středem $\mathbf{x}$ a *nenulovým* poloměrem ε. Množina $P_\varepsilon(\mathbf{x}) = U_\varepsilon(\mathbf{x}) \setminus \{\mathbf{x}\}$ se nazývá **prstencové ε-okolí** bodu $\mathbf{x}$.

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

- **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 $U_\varepsilon(\mathbf{x}) \cap (\mathbb{R}^n \setminus X) \neq \emptyset$.
- **hromadný bod**, jestliže pro každé $\varepsilon > 0$ platí $P_\varepsilon(\mathbf{x}) \cap X \neq \emptyset$.
- **izolovaný bod**, jestliže není hromadný bod.

Všimněte si, že hraniční a hromadný bod množiny nemusí patřit do této množiny. **Vnitřek** [**hranice**] množiny je množina všech jejích vnitřních [hraničních] bodů.

Množina je

- **otevřená**, jestliže všechny její body jsou vnitřní.
- **uzavřená**, jestliže obsahuje každý svůj hraniční bod.

Množina X je uzavřená [otevřená], právě když její doplněk $\mathbb{R}^n \setminus X$ je otevřený [uzavřený]. Otevřenost a uzavřenost se nevylučují: množiny $\emptyset$ a $\mathbb{R}^n$ jsou zároveň otevřené i uzavřené (a jiné takové nejsou). Naopak, některé množiny nejsou otevřené ani uzavřené, např. interval $(0, 1\rangle$.

Množina je **omezená**, jestliže existuje ρ takové, že $\|\mathbf{x} - \mathbf{y}\| < \rho$ pro všechna $\mathbf{x}, \mathbf{y} \in X$. Jinými slovy, X se 'vejde' do koule konečného průměru.

Příklad 6.2. Bod $1/2$ je vnitřním bodem intervalu $\langle 0, 1\rangle$, bod 1 je jeho hraničním bodem. Množina $\{\, (x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1,\ -1 \leq y \leq 1 \,\}$ je uzavřená, množina $\{\, (x, y) \in \mathbb{R}^2 \mid -1 < x < 1,\ -1 < y < 1 \,\}$ je otevřená, bod $(0, 1)$ je hraničním i hromadným bodem obou těchto množin. $\square$

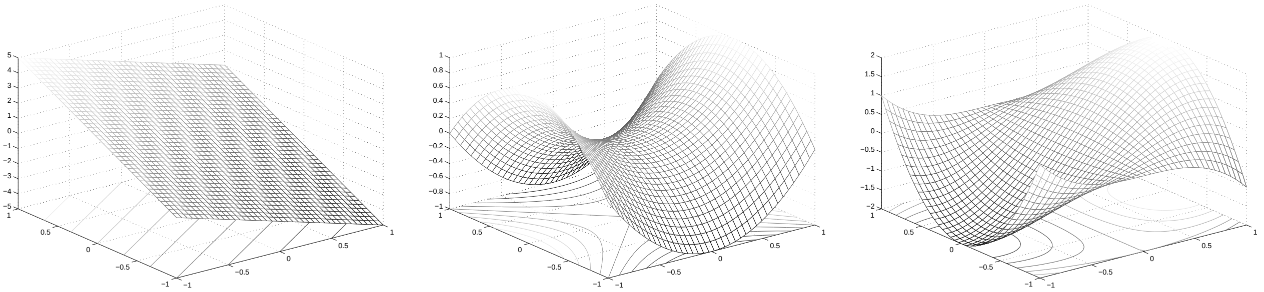
6.3 Zobrazení mezi eukleidovskými prostory

Dále se budeme zabývat zobrazeními, které přiřazují n-rozměrnému vektoru m-rozměrný vektor, tedy zobrazeními $\mathbf{f}\colon X \to \mathbb{R}^m$, kde $X \subseteq \mathbb{R}^n$. Množina X je **definiční obor** zobrazení. Pokud je zobrazení dáno vzorcem a definiční obor není uveden, rozumíme jím největší množinu všech bodů v $\mathbb{R}^n$, kde je vzorec definován.

Pro $m = 1$ jsou hodnotami zobrazení jsou skaláry a proto budeme psát jeho jméno kurzívou, f. Pro $m > 1$ jsou hodnotami zobrazení vektory a jméno budeme psát tučně, $\mathbf{f}$. I když přísně vzato znamenají slova 'funkce' a 'zobrazení' jedno a to samé, je často zvykem pro $m = 1$ mluvit o *funkci* a pro $m > 1$ o *zobrazení*. Někdy se zobrazení do $\mathbb{R}^m$ říká též *vektorová funkce*, ale slovo 'zobrazení' má výhodu, že koresponduje s pojmem 'lineárního zobrazení' z lineární algebry.

Máme tyto speciální případy:

- Pro $n = m = 1$ máme dobře známou funkci jedné proměnné.
- Pro $n = 1$ a $m > 1$ máme vlastně m funkcí jedné proměnné, které jsou složkami zobrazení.



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

Obrázek 6.1: Příklady grafu a vrtevnic funkcí dvou proměnných na obdélníku $\langle -1, 1\rangle^2$, vytvořených příkazem `meshc`.

- Pro $n > 1$ a $m = 1$ máme funkci více proměnných.

Pro funkce (tedy $m = 1$) užíváme tyto pojmy:

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

Tyto definice dobře pochopte! Pojem 'graf funkce' jste dosud možná chápali pouze intuitivně jako 'obrázek'.

Příklad 6.3. Příklady funkcí více proměnných (tedy $m = 1$):

1. $X = \{ (x, y) \in \mathbb{R} \mid x^2 + y^2 \le 1 \}$, $f(x, y) = \sqrt{1 - x^2 - y^2}$
2. $X = \mathbb{R}^2$, $f(x, y) = x^2$
3. $X = \mathbb{R}^n$, $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$ (kde $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ jsou dány)
4. $X = \mathbb{R}^n$, $f(\mathbf{x}) = \mathrm{e}^{-\|\mathbf{x}\|^2}$
5. $X = \mathbb{R}^n$, $f(\mathbf{x}) = x_1$
6. $X = \mathbb{R}^n$, $f(\mathbf{x}) = \max\limits_{i=1}^{n} x_i$ $\square$

Příklad 6.4. Příklady zobrazení (tedy $m > 1$):

1. $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{f}(\mathbf{x}) = \mathbf{x}$ (identické zobrazení neboli identita)
2. $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ (lineární zobrazení) a $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ (afinní zobrazení)
3. $\mathbf{f}\colon \langle 0, 2\pi) \to \mathbb{R}^2$, $\mathbf{f}(t) = (\cos t,\ \sin t)$ (parametrická rovnice kružnice)
4. $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^3$, $\mathbf{f}(t) = (\cos t,\ \sin t,\ at)$ (parametrická rovnice šroubovice neboli helixu)
5. $\mathbf{f}\colon \langle 0, 2\pi)^2 \to \mathbb{R}^3$, $\mathbf{f}(u, v) = (\,(R + r\cos v)\cos u,\ (R + r\cos v)\sin u,\ r\sin v\,)$ (parametrická rovnice toroidu neboli anuloidu)
6. Při technice *image morphing* se obrázek např. obličeje zdeformuje na obrázek jiného obličeje. Pokud množinu bodů obrázku (zidealizovaně) reprezentujeme jako obdélník $X = \langle x_1, x_2\rangle \times \langle y_1, y_2\rangle \subseteq \mathbb{R}^2$, je morphing realizován zobrazením $X \to \mathbb{R}^2$.
7. Elektrické pole přiřadí každému bodu v $\mathbb{R}^3$ vektor z $\mathbb{R}^3$. $\square$

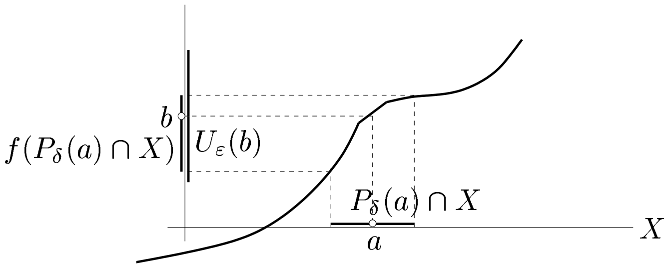
6.4 Limita

Definice 6.1. *Nechť* $\mathbf{a} \in \mathbb{R}^n$ *je hromadný bod množiny* $X \subseteq \mathbb{R}^n$. *Bod* $\mathbf{b} \in \mathbb{R}^m$ *je* **limita** *zobrazení* $\mathbf{f}\colon X \to \mathbb{R}^m$ *v bodě* $\mathbf{a}$, *jestliže pro každé* $\varepsilon > 0$ *existuje* δ *tak, že*

$$\mathbf{f}(P_\delta(\mathbf{a}) \cap X) \subseteq U_\varepsilon(\mathbf{b}). \tag{6.1}$$

Potom značíme

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{b}.$$



Limita popisuje chování zobrazení v blízkosti daného bodu. Blíží-li se vzory k bodu $\mathbf{a}$, obrazy se musí blížit k bodu $\mathbf{b}$. Bod $\mathbf{a}$ nemusí patřit do definičního oboru (protože nezkoumáme obraz přímo bodu $\mathbf{a}$, ale jen bodů v jeho blízkosti), ale musí být jeho hromadným bodem (jinak by nebylo 'dost' bodů blízkých $\mathbf{a}$, které bychom mohli zkoumat).

Mezi limitou funkcí či zobrazení jedné proměnné ($n = 1$) a více proměnných ($n > 1$) je kvalitativní rozdíl daný tím, že množina $\mathbb{R}$ je úplně uspořádaná, kdežto množina $\mathbb{R}^n$ není. Pro $n = 1$ se k bodu $\mathbf{a}$ můžeme blížit pouze dvěma způsoby, zleva a zprava. Aby v tomto bodě existovala limita, musí v něm existovat limita k tomuto bodu zleva a zprava a tyto limity se musí rovnat. To dobře známe z analýzy funkcí jedné proměnné.

Pro $n > 1$ se k bodu $\mathbf{a}$ můžeme blížit nekonečně mnoha způsoby – po různých *množinách*. Každý způsob blížení je dán nějakou množinou $A \subseteq X$ takovou, že $\mathbf{a}$ je jejím hromadným bodem. Aby existovala v bodě $\mathbf{a}$ limita, musí existovat limity odpovídající všem způsobům blížení a musí si být rovny. To lze formalizovat následovně. Označme $\mathbf{f}|_A\colon A \to \mathbb{R}^m$ *restrikci* zobrazení $\mathbf{f}\colon X \to \mathbb{R}^m$ na množinu $A \subseteq X$. Restrikce je to samé zobrazení $\mathbf{f}$ ale s definičním oborem zmenšeným z množiny X na množinu A. Jestliže limita zobrazení $\mathbf{f}$ v bodě $\mathbf{a}$ existuje a je rovna $\mathbf{b}$, pak existuje limita zobrazení $\mathbf{f}|_A$ v bodě $\mathbf{a}$ a je také rovna $\mathbf{b}$.

Příklad 6.5. Pro $m = n = 1$ a $X = \mathbb{R}$ získáme známou limitu funkce jedné proměnné. Volba $A = (-\infty, a)$ vede na limitu v bodě a zleva a volba $A = (a, +\infty)$ vede na limitu v bodě a zprava. Všimněte si, že a je hromadným bodem intervalu A i když do něj nepatří. Aby existovala v bodě a limita, musí existovat limity zleva a zprava a musí si být rovny. □

Příklad 6.6. Nechť $X = \mathbb{R}^2 \setminus (0,0)$ a počítejme

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}\,. \tag{6.2}$$

Položme $A = \{\, (x,y) \in \mathbb{R}^2 \setminus (0,0) \mid y = kx \,\}$, tj. budeme se k bodu $(0,0)$ blížit po přímce jdoucí počátkem se sklonem k. Máme

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{x\to 0} \frac{x \cdot kx}{x^2+(kx)^2} = \frac{k}{1+k^2}.$$

Tato limita je různá pro různá k, tedy limita (6.2) neexistuje. □

Počítat limity pro funkce a zobrazení více proměnných je obtížnější než pro funkce jedné proměnné. Nejsou žádné 'mechanické pomůcky' jako L'Hospitalovo pravidlo, výpočet je často nutné dělat přímo z definice a vyžaduje kreativitu. Naučit se tomuto umění je za rámec našeho kursu.

6.5 Spojitost

Zobrazení $f\colon X \to \mathbb{R}$, kde $X \subseteq \mathbb{R}^n$, je **spojité** v bodě $\mathbf{a} \in X$, jestliže

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}).$$

Zobrazení je spojité na množině X, jestliže je spojité v každém bodě X.

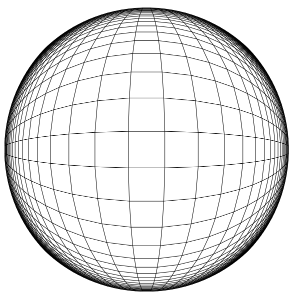
Výše jsme definovali vlastnosti podmnožin $X \subseteq \mathbb{R}^n$ otevřenost, uzavřenost a omezenost. Vyvstává přirozená otázka, které z těchto vlastností se zachovávají spojitým zobrazením. Tedy, zda obraz

$$f(X) = \{\, f(\mathbf{x}) \mid \mathbf{x} \in X \,\}$$

množiny X v zobrazení f má také tyto vlastnosti. Je snadné ukázat, že žádná z těchto vlastností se nezachovává zvlášť.

Příklad 6.7. Nechť $X \subseteq \mathbb{R}$ je interval $\langle 1, +\infty)$. Tato množina je uzavřená a není omezená. Zobrazení $f(x) = 1/x$ (spojité na X) tento interval zobrazí na interval $f(X) = (0, 1\rangle$, který není uzavřený a je omezený. □

Příklad 6.8. Obraz neomezené množiny $X = \mathbb{R}^n$ spojitým zobrazením $\mathbf{f}(\mathbf{x}) = \mathbf{x}/\sqrt{1 + \mathbf{x}^T\mathbf{x}}$ je otevřená omezená množina $\mathbf{f}(X) = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T\mathbf{x} < 1 \,\}$ (jednotková koule bez hranice). Obrázek znázorňuje množinu $\mathbf{f}(X)$, je-li množina X pravidelná mřížka v $\mathbb{R}^2$:



□

Ovšem zachovává se *kombinace* uzavřenosti a omezenosti. Tato kombinace je natolik hluboká vlastnost, že je pro ní zvláštní slovo, nazývá se **kompaktnost**.

Věta 6.1. *Spojité zobrazení uzavřené omezené množiny je uzavřená omezená množina.*

Pro *funkci* $f\colon X \to \mathbb{R}$ je obraz uzavřené omezené množiny $X \subseteq \mathbb{R}^n$ uzavřená omezená podmnožina $\mathbb{R}$. To ale nemůže být nic jiného než uzavřený konečný interval nebo sjednocení takových intervalů. Taková množina jistě má nejmenší a největší prvek. To je důležitý důsledek pro optimalizaci. Je znám jako *věta o extrémní hodnotě* nebo *Weierstrassova věta*.

Důsledek 6.2. *Spojitá funkce $f\colon X \to \mathbb{R}$ nabývá na uzavřené omezené množině $X \subseteq \mathbb{R}^n$ svého minima i svého maxima. Tedy existují prvky $\underline{\mathbf{x}}, \overline{\mathbf{x}} \in X$ takové, že*

$$f(\underline{\mathbf{x}}) = \min f(X) = \min_{\mathbf{x}\in X} f(\mathbf{x}), \qquad f(\overline{\mathbf{x}}) = \max f(X) = \max_{\mathbf{x}\in X} f(\mathbf{x}).$$

6.6 Cvičení

6.1. Pro tyto množiny rozhodněte, zda existují infimum, minimum, supremum, maximum:

a) $\{\, 1/n \mid n \in \mathbb{N} \,\}$
b) $\{\, \mathrm{e}^{-x^2} \mid x \in \mathbb{R} \,\}$
c) $\{\, x \in \mathbb{Q} \mid x^2 \le 2 \,\}$
d) $\{\, |x-y| \mid x \in \langle 0, 1\rangle,\ y \in (1, 2\rangle \,\}$

6.2. Jsou tyto množiny uzavřené, otevřené, omezené? Co je jejich vnitřek a hranice?

a) $\{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1,\ y > 0 \,\} \cup \{1, 1\}$
b) $\{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1,\ y \ge 0 \,\}$
c) $\{\, (x, y) \in \mathbb{R}^2 \mid y = x^2,\ -1 < x \le 1 \,\}$
d) $\{\, (x, y) \in \mathbb{R}^2 \mid xy < 1,\ x > 0,\ y > 0 \,\}$
e) $\mathbb{Q}^3$
f) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \max_{i=1}^n x_i \le 1 \,\}$
g) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} = b \,\}$, kde $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ (nadrovina)
h) $\{\, \mathbf{x} \in \mathbb{R}^n \mid b \le \mathbf{a}^T\mathbf{x} \le c \,\}$, kde $\mathbf{a} \in \mathbb{R}^n$, $b, c \in \mathbb{R}$ (panel)
i) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \,\}$, kde $\mathbf{A}$ je široká (afinní podprostor $\mathbb{R}^n$)
j) $\{\, \mathbf{A} \in \mathbb{R}^{n\times n} \mid \det \mathbf{A} = 0 \,\}$
k) $\{\, \mathbf{A} \in \mathbb{R}^{n\times n} \mid \mathbf{A}^T\mathbf{A} = \mathbf{I} \,\}$

6.3. Může být bod zároveň vnitřní i hraniční? Najděte příklad nebo vyvraťte z definice.

6.4. Matlab má bohatou podporu na vizualizaci funkcí dvou proměnných, včetně kreslení grafu a vrstevnic. Např. pro funkci $f(x, y) = x^2 - y^2$ na obdélníku $\langle -1, 1\rangle^2$ vytvořte data,

```
N=40; x=ones(N,1)*linspace(-1,1,N); y=x'; z=x.^2-y.^2;
```

a pak zkoumejte tyto příkazy: `mesh(z)`, `mesh(x,y,z)`, `meshc(x,y,z)`, `surf(x,y,z)`, `surfc(x,y,z)`, `contour(z)`, `contour(x,y,z)`, `contour(x,y,z,20)`, `imagesc(z)`.

6.5. Co je vrstevnice výšky 1 funkce $f(x, y) = xy$?

6.6. Co jsou vrstevnice funkce $f(\mathbf{x}) = \|\mathbf{x}\|$?

6.7. Co znamená funkce $f(\mathbf{x}) = \inf\{\, \|\mathbf{x} - \mathbf{y}\| \mid \mathbf{y} \in X \,\}$ pro danou množinu $X \subseteq \mathbb{R}^n$? Jak budou vypadat její vrstevnice pro $X = \langle -1, 1\rangle^n$ (načrtněte pro $n = 2$)?

Kapitola 7

Derivace

Předpokládáme, že student zná derivace funkcí jedné proměnné a parciální derivace funkcí více proměnných. Parciální derivaci funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ podle x_i značíme jedním ze symbolů

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

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

Příklad 7.1. Parciální derivace funkce $f(x,y) = x^2y + \sin(x - y^3)$ podle x a podle y jsou

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= f_x(x,y) = 2xy + \cos(x - y^3)\\ \frac{\partial f(x,y)}{\partial y} &= f_y(x,y) = x^2 - 3y^2\cos(x - y^3).\end{aligned}$$

□

V této kapitole dále uvažujeme funkce a zobrazení definované na celém $\mathbb{R}^n$ místo na množině $X \subseteq \mathbb{R}^n$. Zjednoduší to značení a žádnou podstatnou myšlenku tím neztratíme.

7.1 Totální derivace zobrazení

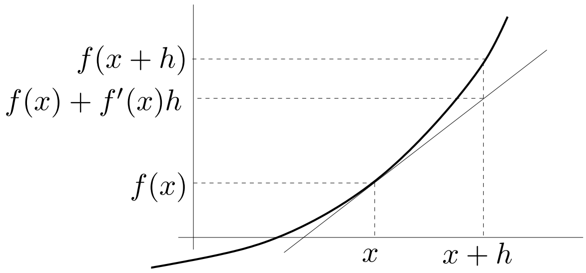
Zopakujme definici derivace funkce jedné proměnné $f\colon \mathbb{R} \to \mathbb{R}$ v bodě x. Existuje-li limita

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}, \tag{7.1}$$

funkce se nazývá diferencovatelná v bodě x a hodnota limity její derivace v bodě x. Diferencovatelnost znamená, že funkci lze v blízkosti bodu x ‘dobře aproximovat’ jako

$$f(x+h) \approx f(x) + f'(x)h. \tag{7.2}$$

Viz obrázek:



Jak zobecnit tuto definici pro zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$? Tvar (7.1) se ukazuje jako k tomu nevhodný, vhodný je ale tvar (7.2). Zkusme zobrazení $\mathbf{f}$ aproximovat v blízkosti bodu $\mathbf{x}$ jako

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{f}'(\mathbf{x})\mathbf{h}. \tag{7.3}$$

Považujeme-li $\mathbf{x}$ za konstantu, pravá strana je afinní zobrazení v proměnné $\mathbf{h}$. Zatím netušíme, zda je taková aproximace vůbec možná a co by mohl znamenat symbol $\mathbf{f}'(\mathbf{x})$. Nicméně je jasné, že $\mathbf{f}'(\mathbf{x})$ musí být matice rozměru $m \times n$. Ukazuje se, že tato matice má zcela přirozený tvar, totiž její složky jsou parciální derivace všech složek zobrazení podle všech proměnných,

$$\frac{\mathrm{d}\mathbf{f}(\mathbf{x})}{\mathrm{d}\mathbf{x}} = \mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix}. \tag{7.4}$$

Tato matice se nazývá **totální derivace**, nebo krátce jen **derivace**. Z historických důvodů se jí říká také **Jacobiho matice**.

Teď je ale nutno celou věc postavit na pevnou zem. Musíme definovat, co znamená, že (7.3) je 'dobrá aproximace'. Chceme, aby chyba aproximace $\mathbf{r}(\mathbf{h}) = \mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}$ byla malá pro malé $\mathbf{h}$. Jak tento požadavek formalizovat? První nápad je požadovat $\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{r}(\mathbf{h}) = \mathbf{0}$, ale to platí vždy. Správná volba je požadovat $\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{r}(\mathbf{h})/\|\mathbf{h}\| = \mathbf{0}$, tedy chyba musí klesat se zmenšujícím se $\mathbf{h}$ 'rychleji než lineárně'.

Definice 7.1. *Zobrazení* $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ *se nazývá* **diferencovatelné** *v bodě* $\mathbf{x}$, *jestliže existuje matice* $\mathbf{f}'(\mathbf{x}) \in \mathbb{R}^{m\times n}$ *taková, že*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}}{\|\mathbf{h}\|} = \mathbf{0}. \tag{7.5}$$

Je-li zobrazení $\mathbf{f}$ diferencovatelné, lze ukázat, že matice $\mathbf{f}'(\mathbf{x})$ je požadavkem (7.5) určena jednoznačně, existují všechny parciální derivace a platí (7.4).

Někdy 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. Rozdíl je přesně stejný, jako když v lineární algebře místo 'lineární zobrazení' říkáme pouze 'matice'.

Příklad 7.2. Ukážeme, že pro $n = m = 1$ totální derivace splývá s obyčejnou derivací (7.1). Požadavek (7.5) v tom případě zní

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - f'(x)h}{|h|} = 0, \tag{7.6}$$

kde $x, h, f(x), f'(x)$ jsou skaláry. Použijeme fakt, že pro $h \geq 0$ je $|h| = h$ a pro $h \leq 0$ je $|h| = -h$. Rovnice (7.6) je tedy ekvivalentní dvěma rovnicím

$$\lim_{h\to 0+} \frac{f(x+h) - f(x) - f'(x)h}{h} = 0, \qquad \lim_{h\to 0-} \frac{f(x+h) - f(x) - f'(x)h}{-h} = 0.$$

Ale ve zlomku $f'(x)h/h$ se h vykrátí a tyto dvě rovnice tedy lze psát jako

$$\lim_{h\to 0+} \frac{f(x+h) - f(x)}{h} = f'(x), \qquad \lim_{h\to 0-} \frac{f(x+h) - f(x)}{-h} = -\lim_{h\to 0-} \frac{f(x+h) - f(x)}{h} = -f'(x).$$

Levá rovnice říká, že $f'(x)$ je derivace funkce f v bodě x zprava. Pravá rovnice říká, že $f'(x)$ je derivace f v x zleva. Tedy rovnice (7.5) požaduje, aby funkce f měla v x derivaci zprava i zleva a obě byly rovny $f'(x)$. $\square$

Pro $n = 1$ je zobrazení $\mathbf{f}$ vektorová funkce jedné proměnné a jeho derivace je *sloupcový vektor* (matice $m \times 1$) sestavený z derivací jednotlivých složek, $\mathbf{f}'(x) = (f_1'(x), \ldots, f_m'(x))$. Není to tedy nic nového oproti derivaci funkcí jedné proměnné. Pro $m = 1$ je zobrazení $\mathbf{f}$ funkce n proměnných $f\colon \mathbb{R}^n \to \mathbb{R}$ a jeho derivací je *řádkový vektor* (matice $1 \times n$).

Jak zjistit, zda je zobrazení diferencovatelné? Zjišťovat to přímo z Definice 7.1 je obtížné a tudíž nepraktické. Je jasné, že pro diferencovatelnost zobrazení nepostačuje jeho spojitost: např. funkce $f(x) = |x|$ je v bodě 0 spojitá ale ne diferencovatelná. Člověk by doufal, že pro diferencovatelnost postačí existence všech parciálních derivací v bodě $\mathbf{x}$. Bohužel, ani to nestačí, neboť parciální derivace hovoří jen o chování funkce na řezech podle souřadnicových os.

Postačující (i když ne nutnou) podmínkou pro diferencovatelnost je, aby všechny parciální derivace *existovaly* a byly to *spojité funkce* proměnné $\mathbf{x}$. Tato podmínka se většinou snadno ověří a vystačíme s ní ve většině praktických situací.

Příklad 7.3. Nechť je funkce $\mathbb{R}^2 \to \mathbb{R}$ 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 tomto bodě sice existují obě parciální derivace (obě jsou rovny nule), ale parciální derivace podle x není v tomto bodě spojitou funkcí (x, y). Lze ukázat z definice, že v bodě $(0, 0)$ funkce není diferencovatelná – což nepřekvapuje, neboť funkce s v okolí tohoto bodu afinní funkci vůbec nepodobá. □

7.2 Směrová derivace

Řez zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{s} \in \mathbb{R}^n$ je zobrazení $\boldsymbol{\varphi}\colon \mathbb{R} \to \mathbb{R}^m$ dané jako

$$\boldsymbol{\varphi}(h) = \mathbf{f}(\mathbf{x} + h\mathbf{s}).$$

Směrovou derivací zobrazení $\mathbf{f}$ v bodě $\mathbf{x}$ ve směru $\mathbf{s}$ nazýváme číslo

$$\boldsymbol{\varphi}'(0) = \left.\frac{\mathrm{d}\boldsymbol{\varphi}(h)}{\mathrm{d}h}\right|_{h=0} = \lim_{h \to 0} \frac{\mathbf{f}(\mathbf{x} + h\mathbf{s}) - \mathbf{f}(\mathbf{x})}{h}. \tag{7.7}$$

Zde $\boldsymbol{\varphi}'(h) = (\varphi_1'(h), \ldots, \varphi_m'(h))$ je derivace zobrazení $\boldsymbol{\varphi}$, kde φ_j jsou funkce jedné proměnné.

Směrová derivace ve směru i-tého vektoru standardní báze $\mathbf{e}_i$ není nic jiného než parciální derivace podle proměnné x_i.

Směrová derivace diferencovatelného zobrazení se dá snadno spočítat z parciálních derivací.

Věta 7.1. *Nechť* $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$, *je diferencovatelné zobrazení. Pak směrová derivace* $\mathbf{f}$ *v bodě* $\mathbf{x}$ *ve směru* $\mathbf{s}$ *je rovna* $\mathbf{f}'(\mathbf{x})\mathbf{s}$.

Důkaz. Chceme dokázat, že výraz (7.7) je roven $\mathbf{f}'(\mathbf{x})\mathbf{s}$. Zkoumejme absolutní hodnotu rozdílu těchto výrazů. Je jasné, že ta je rovna

$$\left| \lim_{h \to 0} \frac{\mathbf{f}(\mathbf{x} + h\mathbf{s}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})h\mathbf{s}}{h} \right| = \lim_{h \to 0} \left| \frac{\mathbf{f}(\mathbf{x} + \mathbf{t}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{t}}{\|\mathbf{t}\|} \right| \|\mathbf{s}\|,$$

kde $\mathbf{t} = h\mathbf{s}$. Ale limita vpravo je restrikce limity (7.5) na řez $\mathbf{t} = h\mathbf{s}$, proto je rovna nule. □

Nic takového jako Věta 7.1 neplatí, když $\mathbf{f}$ není diferencovatelné. Lze ukázat, že zobrazení $\mathbf{f}$ je diferencovatelné v bodě $\mathbf{x}$ právě tehdy, když existuje matice $\mathbf{f}'(\mathbf{x})$ tak, že směrová derivace ve směru $\mathbf{s}$ je rovna $\mathbf{f}'(\mathbf{x})\mathbf{s}$. Tedy existence všech směrových derivací ještě pro diferencovatelnost nepostačuje, ale postačuje, když směrová derivace je lineární zobrazení směru.

7.3 Totální derivace složeného zobrazení

'Řetězové pravidlo' pro derivaci složených funkcí lze přirozeným způsobem rozšířit na zobrazení.

Věta 7.2. *Nechť* $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ *a* $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ *jsou diferencovatelná zobrazení. Derivace složeného zobrazení* $\mathbf{f} \circ \mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^\ell$ *je*

$$(\mathbf{f} \circ \mathbf{g})'(\mathbf{x}) = \frac{\mathrm{d}\mathbf{f}(\mathbf{g}(\mathbf{x}))}{\mathrm{d}\mathbf{x}} = \mathbf{f}'(\mathbf{g}(\mathbf{x}))\,\mathbf{g}'(\mathbf{x}). \tag{7.8}$$

Dimenze zúčastněných prostorů lze přehledně znázornit diagramem

$$\mathbb{R}^n \xrightarrow{\mathbf{g}} \mathbb{R}^m \xrightarrow{\mathbf{f}} \mathbb{R}^\ell. \tag{7.9}$$

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

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{u}}\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}}, \tag{7.10}$$

což se dobře pamatuje, protože d$\mathbf{u}$ se 'vykrátí'. Zdůrazněme, že tato rovnost je *násobení matic*. Výraz na levé straně je matice $\ell \times n$, první výraz na pravé straně je matice $\ell \times m$ a druhý výraz na pravé straně je matice $m \times n$. Pro $\ell = m = n = 1$ dostaneme známé 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 7.4. Nechť $f(u, v)$ je diferencovatelná funkce dvou proměnných. Určeme (totální) derivaci funkce $f(x + y, xy)$ podle vektoru (x, y), neboli její parciální derivace podle x a y.

Máme diagram $\mathbb{R}^2 \xrightarrow{\mathbf{g}} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$, kde zobrazení $\mathbf{g}$ je dané předpisem

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{g}(x, y) = \begin{bmatrix} x + y \\ xy \end{bmatrix}.$$

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

$$f'(u, v) = \begin{bmatrix} \frac{\partial f(u,v)}{\partial u} & \frac{\partial f(u,v)}{\partial v} \end{bmatrix} = \begin{bmatrix} f_u(u, v) & f_v(u, v) \end{bmatrix}.$$

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

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

Derivace zobrazení $f \circ \mathbf{g}\colon \mathbb{R}^2 \to \mathbb{R}$ podle vektoru (x, y) je matice 1×2 (tj. řádkový vektor)

$$\begin{aligned} \frac{\mathrm{d}f(\mathbf{g}(x, y))}{\mathrm{d}(x, y)} &= f'(u, v)\mathbf{g}'(x, y) \\ &= \begin{bmatrix} f_u(u, v) & f_v(u, v) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix} \\ &= \begin{bmatrix} f_u(u, v) + y f_v(u, v) & f_u(u, v) + x f_v(u, v) \end{bmatrix}. \end{aligned}$$

□

Příklad 7.5. Ukažme dva způsoby, jak spočítat parciální derivaci f_x funkce $f(x, y) = \mathrm{e}^{(x+y)^2+(xy)^2}$:

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

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

(b) Položme $u = x + y$, $v = xy$, $f(u, v) = \mathrm{e}^{u^2+v^2}$. Z Příkladu 7.4 máme $f_x = f_u + yf_v$. Jelikož

$$f_u = 2u\mathrm{e}^{u^2+v^2}, \quad f_v = 2v\mathrm{e}^{u^2+v^2},$$

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

Příklad 7.6. Spočítejme derivaci funkce $z = f(t + t^2, \sin t)$ podle t.

Máme diagram $\mathbb{R} \xrightarrow{\mathbf{g}} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$, kde zobrazení $\mathbf{g}$ je dané předpisem

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{g}(t) = \begin{bmatrix} t + t^2 \\ \sin t \end{bmatrix}.$$

Máme

$$\begin{aligned} \frac{\mathrm{d}f(t + t^2, \sin t)}{\mathrm{d}t} &= f'(u, v)\mathbf{g}'(t) \\ &= \begin{bmatrix} f_u(u, v) & f_v(u, v) \end{bmatrix} \begin{bmatrix} 1 + 2t \\ \cos t \end{bmatrix} = f_u(u, v)(1 + 2t) + f_v(u, v)\cos t. \end{aligned}$$ □

Příklad 7.7. Dokažme jiným způsobem Větu 7.1. Zobrazení $\mathbf{y} = \boldsymbol{\varphi}(h) = \mathbf{f}(\mathbf{x}+h\mathbf{s})$ je složením dvou zobrazení $\mathbf{y} = \mathbf{f}(\mathbf{u})$ a $\mathbf{u} = \mathbf{x} + h\mathbf{s}$. Máme $\mathrm{d}\mathbf{u}/\mathrm{d}h = \mathbf{s}$. Podle řetězového pravidla je

$$\boldsymbol{\varphi}'(h) = \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}h} = \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{u}}\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}h} = \frac{\mathrm{d}\mathbf{f}(\mathbf{u})}{\mathrm{d}\mathbf{u}}\mathbf{s}.$$

Pro $h = 0$ je $\mathbf{u} = \mathbf{x}$, čímž dostaneme (7.7). □

7.4 Gradient funkce

Často se používá zvláštní název pro *transpozici* totální derivace funkce $f\colon \mathbb{R}^n \to \mathbb{R}$, říká se jí **gradient** a značí se

$$f'(\mathbf{x})^T = \nabla f(\mathbf{x})$$

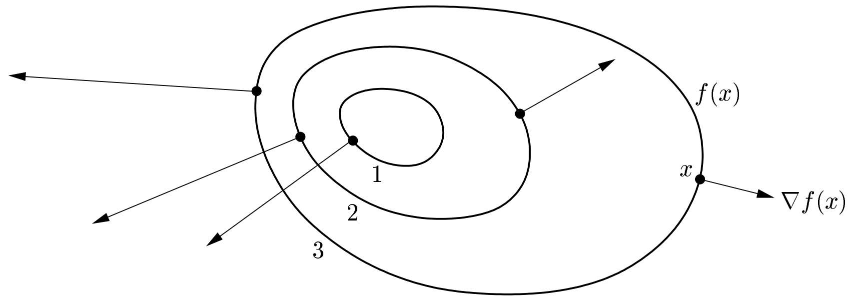
(čteme 'nabla'). Jelikož $f'(\mathbf{x})$ je řádkový vektor, je gradient *sloupcový vektor* parciálních derivací. Zavedení nového slova pro transpozici derivace se zdá zbytečné – nicméně derivace je *lineární funkce*, kdežto gradient je *vektor*.

Literatura ovšem není jednotná v rozlišování gradientu a (totální) derivace. Někdy se obojí ztotožňuje a značí $\nabla f(\mathbf{x})$. To ale vede k nekonzistenci, protože derivace funkce $\mathbb{R}^n \to \mathbb{R}$ pak není speciálním případem derivace zobrazení $\mathbb{R}^n \to \mathbb{R}^m$ (tedy Jacobiho matice) pro $m = 1$, což je řádkový vektor. Dále se budeme symbolu $\nabla f(\mathbf{x})$ spíše vyhýbat a místo něj používat $f'(\mathbf{x})^T$.

Zkoumejme směrovou derivaci $f'(\mathbf{x})\mathbf{s}$ v pevném bodě $\mathbf{x}$ pro různé *normalizované* směry $\mathbf{s}$ (tedy $\|\mathbf{s}\| = 1$). Je jasné, že:

- Směrová derivace je největší ve směru $\mathbf{s} = \nabla f(\mathbf{x})/\|\nabla f(\mathbf{x})\|$, tedy když je $\mathbf{s}$ rovnoběžný s gradientem a stejně orientovaný. Tedy směr gradientu je *směr největšího růstu* funkce.
- Velikost gradientu $\|\nabla f(\mathbf{x})\|$ je velikost strmosti funkce ve směru největšího růstu.
- Směrová derivace ve směru kolmém na gradient je nulová. Jelikož směrová derivace je nulová ve směru vrstevnice, gradient je vždy *kolmý k vrstevnici*.

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



7.5 Parciální derivace druhého řádu

Stejně jako máme obyčejné derivace druhého řádu, máme i parciální derivace druhého řádu. Zderivujeme-li funkci $f\colon \mathbb{R}^n \to \mathbb{R}$ nejdříve podle proměnné x_i a pak podle proměnné x_j, výsledek 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

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

Pokud jsou druhé smíšené parciální derivace

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

spojité funkce v bodě $\mathbf{x}$, pak jsou si rovny. Tedy, pořadí derivování podle jednotlivých proměnných jde zaměnit.

Příklad 7.8. Určeme všechny druhé parciální derivace funkce $f(x,y) = x^2y + \sin(x - y^3)$ z Příkladu 7.1. První derivace už tam máme. Nyní druhé derivace:

$$\begin{aligned}
\frac{\partial^2 f(x,y)}{\partial x^2} &= \frac{\partial}{\partial x}[2xy + \cos(x-y^3)] &&= 2y - \sin(x-y^3)\\
\frac{\partial^2 f(x,y)}{\partial x\,\partial y} &= \frac{\partial}{\partial y}[2xy + \cos(x-y^3)] &&= 2x + 3y^2\sin(x-y^3)\\
\frac{\partial^2 f(x,y)}{\partial y\,\partial x} &= \frac{\partial}{\partial x}[x^2 - 3y^2\cos(x-y^3)] &&= 2x + 3y^2\sin(x-y^3)\\
\frac{\partial^2 f(x,y)}{\partial y^2} &= \frac{\partial}{\partial y}[x^2 - 3y^2\cos(x-y^3)] &&= -6y\cos(x-y^3) - 9y^4\sin(x-y^3).
\end{aligned}$$

□

Pro funkci $f\colon \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$? Nebyla by to již dvojrozměrná tabulka (tedy matice) $n \times n$, nýbrž třírozměrná tabulka velikosti $m \times n \times n$.

7.6 Derivace vektorových a maticových výrazů

Jsou-li funkce nebo zobrazení zadány výrazem obsahujícím vektory a matice, derivaci lze vždy spočítat tak, že výraz rozepíšeme do složek a spočítáme parciální derivace všech složek podle všech proměnných. Pak se snažíme uspořádat tyto parciální derivace do hezké maticové formy. To ale může to být zdlouhavé, proto je dobré si pamatovat derivace často potkávaných výrazů v maticové formě. Např.

$\mathbf{f}(\mathbf{x})$	$\mathbf{f}'(\mathbf{x})$
$\mathbf{x}$	$\mathbf{I}$
$\mathbf{A}\mathbf{x}$	$\mathbf{A}$
$\mathbf{x}^T\mathbf{x}$	$2\mathbf{x}^T$
$\mathbf{x}^T\mathbf{A}\mathbf{x}$	$\mathbf{x}^T(\mathbf{A}+\mathbf{A}^T)$
$\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$	$\mathbf{y}^T$
$\\|\mathbf{x}\\|$	$\mathbf{x}^T/\\|\mathbf{x}\\|$
$\mathbf{g}(\mathbf{A}\mathbf{x})$	$\mathbf{g}'(\mathbf{A}\mathbf{x})\mathbf{A}$
$\mathbf{g}(\mathbf{x})^T\mathbf{g}(\mathbf{x})$	$2\mathbf{g}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$
$\mathbf{g}(\mathbf{x})^T\mathbf{h}(\mathbf{x})$	$\mathbf{g}(\mathbf{x})^T\mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$

Dále je dobré pamatovat si Hessovu matici funkce $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$, tedy kvadratické formy. Je to $f''(\mathbf{x}) = \mathbf{A} + \mathbf{A}^T$.

Odvoďte tyto derivace! Často je pro to užitečné řetězové pravidlo.

7.7 Taylorův polynom

Mějme funkci jedné proměnné, $f\colon \mathbb{R} \to \mathbb{R}$, která má v bodě x_0 derivace až do řádu k. **Taylorův polynom** stupně k je funkce $T_k\colon \mathbb{R} \to \mathbb{R}$ daná předpisem

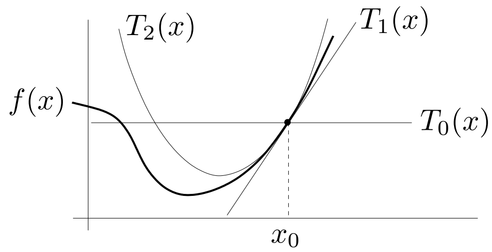
$$T_k(x) = \sum_{i=0}^{k} \frac{1}{i!} f^{(i)}(x_0)\,(x-x_0)^i, \tag{7.11}$$

kde symbol $f^{(i)}$ označuje i-tou derivaci funkce f a kde klademe $0! = 1$. Polynom T_k má zajímavou vlastnost, že v bodě x_0 má všechny derivace až do řádu k (včetně nulté derivace, což je funkční hodnota sama) stejné jako funkce f. Zkuste si to! V tomto smyslu je polynom T_k aproximací funkce f v okolí bodu x_0.

Tvary polynomu až do řádu 2:

$$\begin{aligned}
T_0(x) &= f(x_0)\\
T_1(x) &= f(x_0) + f'(x_0)\,(x-x_0)\\
T_2(x) &= f(x_0) + f'(x_0)\,(x-x_0) + \frac{1}{2}f''(x_0)\,(x-x_0)^2.
\end{aligned}$$

Taylorův polynom nultého řádu $T_0(x)$ je hodně špatná aproximace funkce $f(x)$, rovná jednoduše konstantní funkci $y = f(x_0)$. Polynom prvního řádu $T_1(x)$ není nic jiného než tečna ke grafu funkce f v bodě x_0. Polynom druhého řádu $T_2(x)$ je parabola, která má s funkcí f v bodě x_0 společnou hodnotu a první dvě derivace. Viz obrázek:



Jak zobecnit Taylorův polynom pro funkci více proměnných $f\colon \mathbb{R}^n \to \mathbb{R}$? Nebudeme uvádět obecný vzorec pro celý polynom a uvedeme jen polynom stupně dva. Vyšší stupně se v optimalizaci používají zřídka. Tedy

$$T_1(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0) \tag{7.12a}$$

$$T_2(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T f''(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0). \tag{7.12b}$$

Zde $\mathbf{x}$ a $\mathbf{x}_0$ jsou sloupcové vektory délky n, $f'(\mathbf{x}_0)$ je derivace funkce f (Jacobiho matice, zde řádkový vektor délky n) a $f''(\mathbf{x}_0)$ je Hessova matice (symetrická matice $n\times n$). Ujasněte si, že (7.12a) je afinní funkce $\mathbf{x}$ a (7.12b) je kvadratická funkce $\mathbf{x}$, tj. má tvar (5.8).

Taylorův polynom lze definovat i pro zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ tak, že jednoduše vezmeme Taylorovy polynomy všech složek $f_1,\dots,f_m$. Polynom prvního stupně lze napsat v maticové formě jako

$$\mathbf{T}_1(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0),$$

což není nic jiného než (7.3) po substituci $\mathbf{x}_0 \leftarrow \mathbf{x}$ a $\mathbf{x} \leftarrow \mathbf{x}+\mathbf{h}$. Polynom druhého stupně je zobrazení $\mathbf{T}_2$, jehož složky jsou funkce (7.12b). To nejde napsat v jednoduché maticové formě jako $\mathbf{T}_1$, protože všech $m\times n\times n$ druhých parciálních derivací se 'nevejde' do matice.

7.8 Cvičení

7.1. 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.

7.2. Spočítejte derivaci funkce $g(\mathbf{u}) = f(\mathbf{a}^T\mathbf{u}, \|\mathbf{u}\|)$ podle vektoru $\mathbf{u}$.

7.3. *Metoda konečných diferencí* počítá derivaci funkce přibližně jako $f'(x) \approx [f(x+h) - f(x)]/h$, kde h je malé číslo (dobrá volba je $h = \sqrt{\varepsilon}$, kde ε je strojová přesnost). Toto jde zjevně použít i na parciální derivace. Vymyslete si dvě zobrazení $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ a $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^\ell$ pro nějaké navzájem různé dimenze $n, m, \ell > 1$. 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 $\mathbf{f}'(\mathbf{g}(\mathbf{x}))$.

7.4. 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.

7.5. Spočítejte druhou derivaci $f''(x, y)$ (tj. Hessovu matici) funkcí

a) $f(x, y) = \mathrm{e}^{-x^2 - y^2}$
b) $f(x, y) = \ln(\mathrm{e}^x + \mathrm{e}^y)$

7.6. Je dána funkce $f(x, y) = 6xy^2 - 2x^3 - 3y^3$. V bodě $(x_0, y_0) = (1, -2)$ určete Taylorova polynomu prvního a druhého stupně.

Kapitola 8

Analytické podmínky na lokální extrémy

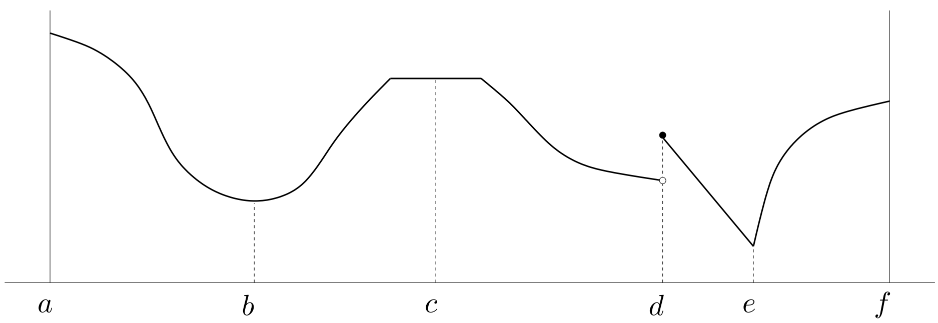
8.1 Globální a lokální extrémy

Funkce $f\colon X \to \mathbb{R}$ má na množině $X \subseteq \mathbb{R}^n$ v bodě $\mathbf{x} \in X$

- **(globální) minimum**, jestliže $f(\mathbf{x}) \le f(\mathbf{y})$ pro všechna $\mathbf{y} \in X$.
- **lokální minimum**, existuje-li $\varepsilon > 0$ tak, že $\mathbf{x}$ je globální minimum funkce f na množině $U_\varepsilon(\mathbf{x}) \cap X$, neboli $f(\mathbf{x}) \le f(\mathbf{y})$ pro všechna $\mathbf{y} \in U_\varepsilon(\mathbf{x}) \cap X$.
- **ostré lokální minimum**, existuje-li $\varepsilon > 0$ tak, že $\mathbf{x}$ je jediné globální minimum funkce f na množině $U_\varepsilon(\mathbf{x}) \cap X$, neboli $f(\mathbf{x}) < f(\mathbf{y})$ pro všechna $\mathbf{y} \in P_\varepsilon(\mathbf{x}) \cap X$.

Podobně se definuje (ostré) lokální maximum. Přívlastek 'globální' se užívá pro zdůraznění, že nejde o minimum lokální. Globální minimum nemusí existovat, neboť množina $f(X) \subseteq \mathbb{R}$ nemusí mít nejmenší prvek. Na druhou stranu, minim může být i více. Každé globální minimum je zároveň lokální, naopak to ale platit nemusí.

Příklad 8.1. Funkce jedné proměnné na obrázku má na uzavřeném intervalu $\langle a, f\rangle$ v bodě a globální a zároveň ostré lokální maximum, v bodě b ostré lokální minimum, v bodě c lokální maximum ale ne ostré, v bodě d ostré lokální maximum, v bodě e globální a zároveň ostré lokální minimum, v bodě f ostré lokální maximum.



□

Pokud bod patří do množiny, musí být buď vnitřním nebo hraničním bodem této množiny. Extrém funkce na množině se tedy nabývá buď ve vnitřním nebo v hraničním bodě množiny. Pokud se nabývá ve vnitřním bodě, jde o **volný extrém**. Pokud se nabývá v hraničním bodě, jde o **vázaný extrém**.

Nalézt globální extrém funkce mnoha proměnných je obecně těžké. Může to totiž vyžadovat nalezení všech lokálních extrémů, a těch může být velký (exponenciální v počtu proměnných) počet. Příklady takových funkcí uvedeme později. Na druhou stranu, nalezení *nějakého* lokálního extrému funkce je obvykle daleko snazší. Dále v této kapitole se budeme zabývat lokálními extrémy diferencovatelných funkcí.

8.2 Volné extrémy

Bod $\mathbf{x}$ z definičního oboru funkce f nazveme

- **stacionární**, pokud v něm je funkce diferencovatelná a má všechny parciální derivace nulové, tedy $f'(\mathbf{x}) = \mathbf{0}^T$.
- **kritický**, pokud je buď stacionární nebo v něm funkce není diferencovatelná.

Poznamenejme, že literatura ne vždy přesně odlišuje tyto dva pojmy.

Věta 8.1. *Nechť $\mathbf{x}$ je vnitřní bod množiny $X \subseteq \mathbb{R}^n$. Nechť $\mathbf{x}$ je lokální extrém funkce $f\colon X \to \mathbb{R}$ na množině X. Pak $\mathbf{x}$ je kritický bod funkce f.*

Příklad 8.2. V Příkladu 8.1 jsou body a, f hraniční (jsou tam lokální extrémy ale ne volné), body b, c stacionární (a tedy i kritické), body d, e kritické ale ne stacionární.

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, neboť implikace je jen jedním směrem.

Funkce $f(\mathbf{x}) = \|\mathbf{x}\|$ má na kouli $\{\, \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq 1 \,\}$ v bodě $\mathbf{0}$ kritický bod ale ne stacionární, je tam lokální minimum. Dále má funkce vázaná lokální maxima ve všech hraničních bodech koule. □

Věta 8.1 říká, že kritické body jsou body 'podezřelé' z volného lokálního extrému. Hledáme-li tedy volné lokální extrémy, najdeme všechny kritické body a o každém rozhodneme, zda je to lokální extrém či ne. V tom nám pomohou následující podmínky druhého řádu.

Věta 8.2. *Nechť $\mathbf{x}$ je vnitřní bod množiny $X \subseteq \mathbb{R}^n$. Nechť je funkce $f\colon X \to \mathbb{R}$ v bodě $\mathbf{x}$ dvakrát diferencovatelná a nechť $\mathbf{x}$ je stacionární bod.*

- *f má v bodě $\mathbf{x}$ ostré lokální minimum [maximum] na X právě tehdy, když Hessova matice druhých derivací $f''(\mathbf{x})$ je pozitivně [negativně] definitní.*
- *Je-li $f''(\mathbf{x})$ indefinitní, nemá f v $\mathbf{x}$ lokální minimum ani lokální maximum na X.*

Věta nic neříká o případu, kdy je $f''(\mathbf{x})$ pozitivně nebo negativně semidefinitní – pak v bodě $\mathbf{x}$ může nebo nemusí být lokální extrém (např. vezměte funkce $f(x) = x^3$ a $f(x) = x^4$ v bodě $x = 0$). Bod $\mathbf{x}$, ve kterém je $f'(\mathbf{x}) = \mathbf{0}^T$ a matice $f''(\mathbf{x})$ je indefinitní, se nazývá **sedlový bod**.

Příklad 8.3. Vraťme se k přibližnému řešení přeurčené nehomogenní lineární soustavy ve smyslu nejmenších čtverců. Řešení úlohy (4.2) lze získat bez geometrické úvahy, položením derivací rovných nule. Upravíme

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

Podmínka na stacionární bod:

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}(\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{b}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{b}) = 2\mathbf{x}^T\mathbf{A}^T\mathbf{A} - 2\mathbf{b}^T\mathbf{A} = \mathbf{0}^T.$$

Po transpozici dostaneme normální rovnice (4.3).

Druhá derivace (Hessián) účelové funkce je $2\mathbf{A}^T\mathbf{A}$. Pokud $\mathbf{A}$ má plnou hodnost, Hessián je pozitivně definitní a tedy nalezený extrém je lokální minimum. Pokud $\mathbf{A}$ nemá plnou hodnost, Hessián je pozitivně semidefinitní a druh extrému nemůžeme z Věty 8.2 určit. □

8.3 Extrémy vázané rovnostmi

Hledejme extrémy funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ na množině $X = \{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = 0 \,\}$, kde $g\colon \mathbb{R}^n \to \mathbb{R}$. Mluvíme o extrémech funkce f vázaných podmínkou $g(\mathbf{x}) = 0$.

V některých příkladech můžeme všechna řešení rovnice $g(\mathbf{x}) = 0$ vyjádřit explicitně. Pak lze úlohu prostým dosazením převést na úlohu bez omezení a použít Věty 8.1 a 8.2.

Příklad 8.4. Hledejme strany kvádru s jednotkovým objemem a minimálním povrchem. Tedy minimalizujeme funkci $f(x, y, z) = xy+xz+yz$ za podmínky $g(x, y, z) = 1-xyz = 0$. Vyjádříme z podmínky $z = 1/(xy)$, dosadíme to do účelové funkce a použijeme Větu 8.1. Dostaneme jediné lokální minimum $(x, y, z) = (1, 1, 1)$. Dopočítejte! □

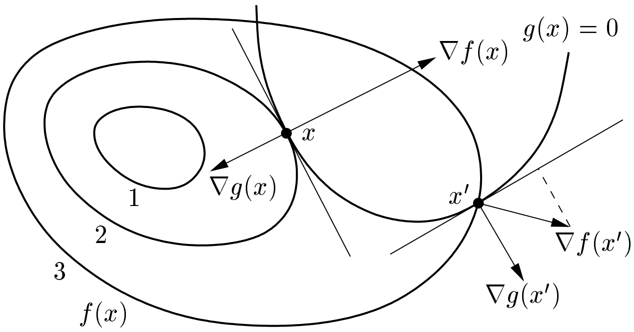
Pokud ale rovnice $g(x) = 0$ explicitně řešit nejde, Věty 8.1 použít obvykle nelze, neboť množina X obvykle nemá žádné vnitřní body. Dále proto popíšeme obecnější přístup, metodu Lagrangeových multiplikátorů.

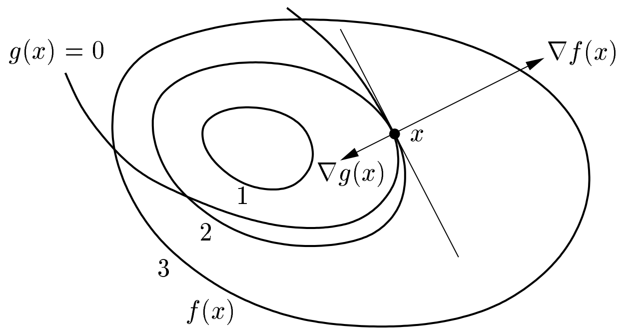
8.3.1 Odvození geometrickou úvahou

Zde odvodíme metodu pomocí jednoduché geometrické úvahy. Zdůrazněme, že tato úvaha nenahrazuje formální důkaz.

Obrázek ukazuje příklad pro $n = 2$. Jsou na něm tři vrstevnice funkce f s výškami 1, 2, 3 a jedna vrstevnice funkce g s výškou 0. Hledá se lokální extrém funkce f na křivce $g(\mathbf{x}) = 0$. Uvažujme bod $\mathbf{x}$ na křivce $g(\mathbf{x}) = 0$ a tečnu ke křivce v tomto bodě. Pokud gradient funkce f v bodě $\mathbf{x}$ má nenulový průmět do tečny, pohyb po křivce ve směru průmětu funkci f určitě zvětší a pohyb proti směru průmětu funkci f určitě zmenší. Bod $\mathbf{x}'$ na levém obrázku tedy nemůže být lokální extrém.

Nulovost průmětu gradientu do tečny ale lokální extrém nezaručuje, je to podmínka nutná nikoli postačující. Bod $\mathbf{x}$ na levém obrázku je lokální minimum: jeho pohyb po křivce $g(\mathbf{x}) = 0$ nahoru i dolů funkci f zvětší. Ovšem bod $\mathbf{x}$ v pravém obrázku není lokální extrém: jeho pohyb po křivce nahoru funkci f zvětší a pohyb dolů ji zmenší. Je to proto, že vrstevnice obou funkcí se v bodě $\mathbf{x}$ křižují.





Došli jsme k následujícímu pozorování: jestliže je bod $\mathbf{x}$ lokální extrém, pak vektor $\nabla f(\mathbf{x})$ má nulový průmět do tečny k vrstevnici ke křivce $g(\mathbf{x}) = 0$ v bodě $\mathbf{x}$. Protože gradient a vrstevnice funkce jsou v každém bodě na sebe kolmé, znamená to rovnoběžnost gradientů funkcí f a g v bodě $\mathbf{x}$. Neboli existuje číslo $\lambda \in \mathbb{R}$ tak, že $\nabla f(\mathbf{x}) = -\lambda \nabla g(\mathbf{x})$, tedy

$$f'(\mathbf{x}) + \lambda\, g'(\mathbf{x}) = \mathbf{0}^T. \tag{8.1}$$

Číslo λ může být kladné i záporné, protože nevíme, na jaké straně vrstevnice $g(\mathbf{x}) = 0$ funkce g klesá a na jaké stoupá. Může být i nulové, což se stane tehdy, když v bodě $\mathbf{x}$ bude gradient f nulový, tedy $\mathbf{x}$ bude stacionární bod funkce f i bez omezení.

Číslo λ se nazývá **Lagrangeův multiplikátor**. Podmínka (8.1) se často zapisuje následujícím ekvivalentním způsobem. Definujeme **Lagrangeovu funkci**

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda\, g(\mathbf{x}). \tag{8.2}$$

Pak podmínka $\partial L(\mathbf{x}, \lambda)/\partial \mathbf{x} = \mathbf{0}^T$ odpovídá rovnici (8.1) a $\partial L(\mathbf{x}, \lambda)/\partial \lambda = g(\mathbf{x}) = 0$ odpovídá omezení. Je-li tedy dvojice $(\mathbf{x}, \lambda)$ stacionární bod Lagrangeovy funkce, je bod $\mathbf{x}$ 'podezřelý' z lokálního extrému funkce f vázaného podmínkou $g(\mathbf{x}) = 0$.

Příklad 8.5. Hledejme extrémy funkce $f(x, y) = x + y$ za podmínky $g(x, y) = 1 - x^2 - y^2 = 0$.

Množina přípustných řešení $X = \{\, (x, y) \mid x^2 + y^2 = 1 \,\}$ je kružnice. 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, $(x, y) = \pm(1, 1)/\sqrt{2}$. U této jednoduché úlohy je jasné, že jde o lokální minimum a maximum a že jsou zároveň globální. □

Úvaha se snadno zobecní na případ $n \geq 2$ tak, že místo tečny ke křivce $g(\mathbf{x}) = 0$ uvažujeme tečnou nadrovinu k nadploše $g(\mathbf{x}) = 0$.

Příklad 8.6. Řešme znovu Příklad 8.4. Lagrangeova funkce je

$$L(x, y, z, \lambda) = xy + xz + yz + \lambda(1 - xyz).$$

Řešíme soustavu

$$\begin{aligned}
\partial L(x, y, z, \lambda)/\partial x &= y + z - \lambda yz &&= 0\\
\partial L(x, y, z, \lambda)/\partial y &= x + z - \lambda xz &&= 0\\
\partial L(x, y, z, \lambda)/\partial z &= x + y - \lambda xy &&= 0\\
\partial L(x, y, z, \lambda)/\partial \lambda &= xyz - 1 &&= 0.
\end{aligned}$$

Soustava je zjevně splněna pro $(x, y, z, \lambda) = (1, 1, 1, 2)$. □

8.3.2 Kdy metoda nefunguje

V naší úvaze jsme mlčky předpokládali, že gradient $\nabla g(\mathbf{x})$ v bodě $\mathbf{x}$ na křivce $g(\mathbf{x}) = 0$ je nenulový. Pokud by tomu tak nebylo, nastal by problém: rovnice (8.1) by nemusela platit pro žádné λ a přesto by $\mathbf{x}$ mohl být lokální extrém. Tedy tento lokální extrém bychom neodhalili.

Zde je nutno se smířit s myšlenkou, že může existovat funkce g, jejíž vrstevnice $g(\mathbf{x}) = 0$ je 'dobře vychovaná' křivka, ale gradienty funkce g v některých bodech této křivky jsou nulové.

Příklad 8.7. Řešme opět Příklad 8.5, kde ale omezení změníme na $g(x, y) = (1-x^2-y^2)^2 = 0$.

Je jasné, že staré i nové omezení definují stejnou kružnici,

$$X = \{\, (x, y) \in \mathbb{R}^2 \mid 1 - x^2 - y^2 = 0 \,\} = \{\, (x, y) \in \mathbb{R}^2 \mid (1 - x^2 - y^2)^2 = 0 \,\}.$$

Úlohy mají tedy stejné řešení. Gradient nové funkce g je

$$\nabla g(x, y) = g'(x, y)^T = \begin{bmatrix} -4x(1 - x^2 - y^2) \\ -4y(1 - x^2 - y^2) \end{bmatrix}.$$

Vidíme, že $\nabla g(x, y) = \mathbf{0}$ pro všechny body (x, y) splňující $g(x, y) = 0$. Čekáme tedy problém.

Stacionární body Lagrangeovy funkce $L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2)^2$ musí splňovat

$$\begin{aligned} \partial L(x, y, \lambda)/\partial x &= 1 - 4\lambda x(1 - x^2 - y^2) = 0 \\ \partial L(x, y, \lambda)/\partial y &= 1 - 4\lambda y(1 - x^2 - y^2) = 0 \\ \partial L(x, y, \lambda)/\partial \lambda &= (1 - x^2 - y^2)^2 \qquad\quad = 0. \end{aligned}$$

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. □

8.3.3 Obecná formulace

Dovolme nyní více než jedno omezení, tedy

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

kde $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$. Množina X obsahuje řešení soustavy rovnic $g_i(\mathbf{x}) = 0$, $i = 1, \ldots, m$, kde funkce $g_1, \ldots, g_m\colon \mathbb{R}^n \to \mathbb{R}$ jsou složky zobrazení $\mathbf{g}$. Geometrická úvaha pro $m > 1$ je poměrně složitá a proto pouze bez důkazu uvedeme následující větu.

Bod $\mathbf{x}$ nazveme **regulární bod** diferencovatelného zobrazení $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$, má-li Jacobiho matice $\mathbf{g}'(\mathbf{x})$ hodnost m, tj. jsou-li gradienty $\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})$ lineárně nezávislé. Všimněte si, že pro $m = 1$ to znamená jednoduše podmínku $\nabla g(\mathbf{x}) \neq \mathbf{0}$ z §8.3.2.

Věta 8.3. *Mějme funkci $f\colon \mathbb{R}^n \to \mathbb{R}$ a zobrazení $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$. Nechť $\mathbf{x}$ je lokální extrém funkce f vázaný podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$. Nechť jsou f a $\mathbf{g}$ v bodě $\mathbf{x}$ diferencovatelné. Nechť $\mathbf{x}$ je regulární bod zobrazení $\mathbf{g}$. Pak existuje vektor $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ tak, že $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod Lagrangeovy funkce*

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i\, g_i(\mathbf{x}). \tag{8.3}$$

Stacionární body $(\mathbf{x}, \boldsymbol{\lambda})$ funkce L získáme řešením soustavy

$$\begin{aligned} \partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \mathbf{x} &= f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x}) = \mathbf{0}^T \\ \partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \boldsymbol{\lambda} &= \mathbf{g}(\mathbf{x})^T \qquad\qquad\quad = \mathbf{0}^T. \end{aligned}$$

Jako cvičení si ujasněte, proč $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial \boldsymbol{\lambda} = \mathbf{g}(\mathbf{x})^T$.

8.3.4 Jak (ne)určovat druh vázaného etrému

Nechť $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod Lagrangeovy funkce. Pak tedy bod $\mathbf{x}$ je 'podezřelý' z lokálního extrému funkce f vázaného podmínkou $\mathbf{g}(\mathbf{x}) = \mathbf{0}$. Jak poznáme, zda tento podezřelý bod $\mathbf{x}$ je vázané lokální minimum, vázané lokální maximum, či nic z toho? K tomu existují podmínky druhého řádu analogické Větě 8.2, jsou ale dosti složité a proto je vynecháme.

Řekneme pouze, jak druh vázaného extrému v podezřelém bodě $\mathbf{x}$ zjistit *nejde*. Nejde to zjistit podle definitnosti Hessovy matice $L''(\mathbf{x}, \boldsymbol{\lambda})$, tedy použít Větu 8.2 na funkci L. Důvodem je, že pokud $\mathbf{x}$ je vázaný lokální extrém funkce f a $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod funkce L, pak $(\mathbf{x}, \boldsymbol{\lambda})$ obecně není lokální extrém funkce L. Naopak, lze ukázat, že *vždy* je to její sedlový bod.

8.3.5 Podurčené lineární soustavy

V §4 jsme se věnovali přeurčené nehomogenní lineární soustavě $\mathbf{Ax} = \mathbf{b}$ pro $\mathbf{A} \in \mathbb{R}^{m\times n}$. Nyní uvažujme případ, kdy soustava má více než jedno řešení, tedy množina řešení je afinní podprostor $\mathbb{R}^n$. Taková soustava je **podurčená**. To nastane typicky pro $m < n$ (tedy rovnic je méně než neznámých) – tato podmínka však není ani nutná ani postačující.

Z množiny řešení přeurčené soustavy je často užitečné vybrat jediné podle nějakého kritéria. Nejjednodušší kritérium je minimalizovat normu řešení, což vede na úlohu

$$\min\{\, \|\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{Ax} = \mathbf{b}\,\}. \tag{8.4}$$

Tato úloha je známa jako řešení soustavy **ve smyslu nejmenší normy** (*least norm solution*). Podotkněme, že někdy je nutno použít jiná kritéria, viz Cvičení 8.20.

Úlohu (8.4) vyřešíme pomocí Lagrangeových multiplikátorů. Místo normy budeme minimalizovat její čtverec $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$. Lagrangeova funkce je

$$L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{x}^T\mathbf{x} + 2\boldsymbol{\lambda}^T(\mathbf{b} - \mathbf{Ax}),$$

kde přidaná dvojka nemění úlohu (proč?) a usnadní odvození. Je $\partial L(\mathbf{x}, \boldsymbol{\lambda})/\partial\mathbf{x} = 2\mathbf{x}^T - 2\boldsymbol{\lambda}^T\mathbf{A}$ (odvoďte!). Stacionární body $(\mathbf{x}, \boldsymbol{\lambda})$ funkce L tedy získáme řešením soustavy

$$\mathbf{x} = \mathbf{A}^T\boldsymbol{\lambda}, \quad \mathbf{Ax} = \mathbf{b}.$$

Dosadíme $\mathbf{x}$ do druhé rovnice, $\mathbf{AA}^T\boldsymbol{\lambda} = \mathbf{b}$. Pokud matice $\mathbf{A}$ má plnou hodnost (tedy m), máme odtud $\boldsymbol{\lambda} = (\mathbf{AA}^T)^{-1}\mathbf{b}$. Dosazením do první rovnice dostaneme

$$\mathbf{x} = \mathbf{A}^+\mathbf{b}, \quad \text{kde} \quad \mathbf{A}^+ = \mathbf{A}^T(\mathbf{AA}^T)^{-1}. \tag{8.5}$$

Matice $\mathbf{A}^+$ se zde nazývá **pseudoinverze**. Je to jedna z pravých inverzí matice $\mathbf{A}$ (ověřte!).

Pseudoinverzi jsme již jednou definovali vzorcem (4.4). Má-li $\mathbf{A}$ plnou hodnost, její pseudoinverzi je rovna buď $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ nebo $\mathbf{A}^T(\mathbf{AA}^T)^{-1}$ podle toho, zda je $\mathbf{A}$ úzká nebo široká. Vektor $\mathbf{x} = \mathbf{A}^+\mathbf{b}$ je v prvním případě řešení přeurčené soustavy ve smyslu nejmenších čtverců, ve druhém případě řešení podurčené soustavy ve smyslu nejmenší normy.

V případě, že $\mathbf{A}$ nemá plnou hodnost, je řešení složitější.

8.4 Cvičení

8.1. 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

a) $\{2, 3, -1\}$
b) $\{2, 3, 0\}$
c) $\{0, -1, 1\}$

8.2. 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.

a) $f(x, y) = x(1 - \frac{2}{3}x^2 - y^2)$
b) $f(x, y) = 1/x + 1/y + xy$
c) $f(x, y) = \mathrm{e}^y(y^2 - x^2)$

Dávejte dobrý pozor při řešení soustav rovnic vzniklých z podmínky na stacionární bod! Snadno se totiž stane, že vám nějaké řešení unikne.

8.3. Je dána funkce $f\colon \mathbb{R}^n \to \mathbb{R}$, množiny $Y \subseteq X \subseteq \mathbb{R}^n$, a bod $\mathbf{x} \in Y$. Uvažujme dva výroky:

a) Funkce f má v bodě $\mathbf{x}$ lokální minimum na množině X.
b) Funkce f má v bodě $\mathbf{x}$ lokální minimum na množině Y.

Vyplývá druhý výrok z prvního? Vyplývá první výrok z druhého? Dokažte z definice lokálního extrému nebo vyvraťe nalezením protipříkadu.

V následujících úlohách na Lagrangeovy multiplikátory nemusíte určovat druh extrému, stačí pouze najít 'podezřelé' body.

8.4. Najděte lokální extrémy funkcí

a) $f(x, y) = 2x - y$
b) $f(x, y) = x^2 + 2y^2$
c) $f(x, y) = x^2 y$
d) $f(x, y) = x^4 + y^2$
e) $f(x, y) = \sin(xy)$
f) $f(x, y) = \mathrm{e}^{xy}$

na kružnici $x^2 + y^2 = 1$.

Nápověda: Někdy je dobré účelovou funkci zjednodušit, pokud to nezmění řešení. A buďte opět velmi pozorní při řešení vzniklých soustav rovnic.

8.5. Rozložte dané kladné reálné číslo na součin n kladných reálných čísel tak, aby jejich součet byl co nejmenší.

8.6. Pastevec 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). Jaký bude obsah ohrady?

8.7. 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) kelímek (komolý kužel bez víka). Objem komolého kužele a povrch jeho pláště je dán vzorci $V = \frac{\pi}{3}h(R^2 + Rr + r^2)$ a $S = \pi(R+r)\sqrt{(R-r)^2 + h^2}$.

8.8. Najděte bod nejblíže počátku na křivce

a) $x + y = 1$
b) $x + 2y = 5$
c) $y = x^3 + 1$
d) $x^2 + 2y^2 = 1$

8.9. Nechť $\mathbf{x}^*$ je bod nejblíže počátku na nadploše $h(\mathbf{x}) = 0$. Ukažte metodou Lagrangeových multiplikátorů, že vektor $\mathbf{x}^*$ je kolmý k tečné nadrovině plochy v bodě $\mathbf{x}^*$.

8.10. 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 $xy + yz + zx = 1$

8.11. Máme kouli o poloměru r a středu $\mathbf{x}_0$, tj. množinu $\{\, \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| \le r \,\}$. Máme nadrovinu $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} = b \,\}$. Najděte vzdálenost koule od nadroviny metodou Lagrangeových multiplikátorů.

8.12. Do elipsy o daných délkách os vepište obdélník s maximálním obsahem.

8.13. *Fermatův princip* v paprskové optice říká, že cesta mezi libovolnými dvěma body na paprsku má takový tvar, aby ji světlo proběhlo za čas kratší než jí blízké dráhy. Později se zjistilo, že správným kritériem není *nejkratší* ale *extrémní* čas. Tedy skutečná dráha paprsku musí mít čas větší nebo menší než jí blízké dráhy. Z tohoto principu odvoďte:

a) Zákon odrazu od zrcadla: úhel dopadu se rovná úhlu odrazu.
b) Snellův zákon lomu: na rozhraní dvou prostředí se světlo lomí tak, že

$$\frac{c_1}{c_2} = \frac{\sin \alpha_1}{\sin \alpha_2},$$

kde α_i je úhel paprsku od normály rozhraní a c_i je rychlost světla v prostředí i.

Odvození udělejte

a) pro rovinné zrcadlo a rozhraní (což vede na minimalizaci bez omezení),
b) ($\star$) 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?

8.14. Rozdělení pravděpodobnosti diskrétní náhodné proměnné je funkce $p\colon \{1, \dots, n\} \to \mathbb{R}_+$ (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. Udělejte totéž za omezení, že je předepsána střední hodnota $\mu = \sum_{i=1}^{m} x\, p(x)$.
b) Dokažte *Gibbsovu nerovnost* (též zvanou *informační nerovnost*): pro každé dvě rozdělení p, q platí $\sum_{x=1}^{n} p(x) \log q(x) \ge \sum_{x=1}^{n} p(x) \log p(x)$, přičemž rovnost nastává jen tehdy, když $p = q$.

8.15. ($\star$) 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?

8.16. 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.

8.17. Minimalizujte $\mathbf{x}^T\mathbf{x}$ za podmínky $\mathbf{a}^T\mathbf{x} = 1$. Jaký je geometrický význam úlohy?

8.18. Maximalizujte $\mathbf{a}^T\mathbf{x}$ za podmínky $\mathbf{x}^T\mathbf{x} = 1$. Jaký je geometrický význam úlohy?

8.19. Minimalizujte $\mathbf{x}^T\mathbf{A}\mathbf{x}$ za podmínky $\mathbf{b}^T\mathbf{x} = 1$, kde $\mathbf{A}$ je pozitivně definitní.

8.20. Minimalizujte $\|\mathbf{C}\mathbf{x}\|$ za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde $\mathbf{C}$ je čtvercová nebo úzká s lineárně nezávislými sloupci.

8.21. ($\star$) Minimalizujte $\|\mathbf{C}\mathbf{x}\|$ za podmínek $\mathbf{A}\mathbf{x} = \mathbf{0}$ a $\mathbf{x}^T\mathbf{x} = 1$.

8.22. ($\star$) Minimalizujte $\|\mathbf{A}\mathbf{x}\|$ za podmínek $\mathbf{x}^T\mathbf{C}\mathbf{x} = 1$, kde $\mathbf{C}$ je positivně definitní.

8.23. ($\star$) Jaké musí být vlastnosti matice $\mathbf{A}$ a vektoru $\mathbf{b}$, aby $\max\{\, \|\mathbf{A}\mathbf{x}\| \mid \mathbf{b}^T\mathbf{x} = 0 \,\} = 0$?

Kapitola 9

Numerické algoritmy na hledání volných lokálních extrémů

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

9.1 Rychlost konvergence iteračních algoritmů

Iterační algoritmus konstruuje posloupnost bodů $\mathbf{x}_k$, která se blíží k řešení $\mathbf{x}$. Zkoumejme rychlost, s jakou se posloupnost zbytků $a_k = \|\mathbf{x}_k - \mathbf{x}\|$ blíží nule. Tato posloupnost je zjevně nezáporná, $a_k \geq 0$, a pokud algoritmus konverguje, máme $\lim_{k\to\infty} a_k = 0$.

Pokud existuje limita

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = r, \tag{9.1}$$

řekneme, že posloupnost $\{a_k\}$ konverguje

- **sublineárně**, pokud $r = 1$
- **lineárně**, pokud $0 < r < 1$
- **superlineárně**, pokud $r = 0$.

Je jasné, že čím je r menší, tím posloupnost konverguje ‘rychleji’. Sublineární konvergence znamená velmi (často nepoužitelně) pomalý algoritmus. Lineární konvergence znamená přijatelnou rychlost, přibližně rovnou rychlosti konvergence geometrické řady. Většina numerických algoritmů konverguje lineárně. Superlineární konvergence znamená výborný algoritmus.

Příklad 9.1.

1. Posloupnost $\{a_k\} = \{2^{-k}\} = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots\right\}$ konverguje lineárně, protože $a_{k+1}/a_k = 1/2$, což je nezávislé na k. Posloupnost je obyčejná geometrická řada.
2. Posloupnost $\{a_k\} = \{1/k\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right\}$ konverguje sublineárně, protože $a_{k+1}/a_k = k/(k+1)$, což pro $k \to \infty$ se blíží 1.
3. Posloupnost $\{a_k\} = \{2^{-2^k}\} = \left\{\frac{1}{4}, \frac{1}{16}, \frac{1}{256}, \ldots\right\}$ konverguje superlineárně, protože

   $$\frac{a_{k+1}}{a_k} = \frac{2^{-2^{k+1}}}{2^{-2^k}} = 2^{-2^{k+1}+2^k} = 2^{-2^k}$$

   a tedy limita (9.1) je rovna 0.

$\square$

Uvědomte si, jak fantasticky rychlá je to konvergence. Znamená to, že $a_{k+1} = a_k{}^2$, tj. s každou iterací se zhruba *zdvojnásobí počet platných cifer*. Strojové přesnosti dosáhneme za několik málo iterací.

4. Posloupnost $\{a_k\} = \{k^{-k}\}$ konverguje superlineárně (limitu (9.1) spočtěte!).
5. Pro posloupnost $\{a_k\} = \left\{\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \ldots\right\}$, tj. 'koktavou' verzi posloupnosti $\{2^{-k}\}$, limita (9.1) neexistuje, protože výraz a_{k+1}/a_k je jiný pro sudé a pro liché k.

Poslední příklad ukazuje nedostatečnost stávající definice: limita (9.1) neexistuje, přesto že posloupnost je jinak 'rozumná'. Proto se zavádí obecnější definice. Řekneme, že posloupnost $\{a_k\}$ konverguje **alespoň** sublineárně [lineárně, superlinárně], existuje-li posloupnost $\{a'_k\}$, která konverguje sublineárně [lineárně, superlinárně] a $a'_k \geq a_k$ pro každé k. Např. posloupnost z příkladu 5 výše konverguje alespoň lineárně, protože můžeme zvolit $a'_k = 2^{-k/2}$.

9.2 Metoda zlatého řezu

(Tato část je nepovinná.)

Půlení intervalu je známá iterační metoda na hledání nulové hodnoty spojité funkce $g\colon \mathbb{R} \to \mathbb{R}$ (tj. hledání kořene rovnice $g(x) = 0$), bez nutnosti počítání jejích derivací (které ani nemusejí existovat). Na začátku iterace jsou dány dva body x_1 a x_2 takové, že $g(x_1)g(x_2) < 0$. Tedy v intervalu $\langle x_1, x_2\rangle$ musí ležet aspoň jeden kořen. Zjistíme hodnotu funkce $g(x_3)$ v bodě $x_3 = (x_1 + x_2)/2$. 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 $\langle x_1, x_2\rangle$ nahradíme intervalem $\langle x_1, x_3\rangle$, ve druhém případě intervalem $\langle x_3, x_2\rangle$. Pokračujeme stejně. Všimněme si, že v každé iteraci se interval neučitosti zúží na polovinu.

Hledejme nyní nikoliv nulovou hodnotu, ale minimum funkce $f\colon \mathbb{R} \to \mathbb{R}$. Nemáme opět přístup k derivacím (které nemusejí existovat). Přepodkládáme, že funkce je na daném počátečním intervalu $\langle x_1, x_2\rangle$ spojitá a **unimodální** – tj. existuje bod x^* tak, že na intervalu $\langle x_1, x^*\rangle$ funkce striktně klesá a na intervalu $\langle x^*, x_2\rangle$ striktně roste. Zatímco v metodě půlení intervalu stačilo pro zúžení intervalu přidat jediný bod, zde potřebujeme do intervalu umístit dva body, x_3 a x_4. Tedy máme čtyři body tak, že $x_1 < x_3 < x_4 < x_2$. Musí nastat jeden z těchto případů:

1. $f(x_3) \leq f(x_4)$. Minimum je nutně v intervalu $\langle x_1, x_4\rangle$.
2. $f(x_3) > f(x_4)$. Minimum je nutně v intervalu $\langle x_3, x_2\rangle$.

Zůstává otázka, jak volit pozici bodů, aby bylo zaručeno největší možné zmenšení intervalu neurčitosti, a to při obou možnostech 1 a 2. Jedna možnost je zvolit x_3 a x_4 blízko vedle sebe symetricky uprostřed intervalu $\langle x_1, x_2\rangle$, tj.

$$x_3 = \frac{x_1 + x_2}{2} - \delta, \quad x_4 = \frac{x_1 + x_2}{2} + \delta$$

pro nějaké velmi malé $\delta > 0$. Toto však nezaručuje největší možné zúžení intervalu neurčitosti.

Označme $a = x_3 - x_1$, $b = x_2 - x_3$, $c = x_4 - x_3$. Aby byl interval v $(k+1)$-ní iteraci rozdělen ve stejném poměru jako v k-té iteraci, a to nezávisle na tom jaká z možností 1 a 2 nastane, potřebujeme

$$\frac{b}{a} = \frac{a}{c} = \frac{b}{b-c}.$$

Odtud dostaneme $\varphi - \varphi^{-1} = 1$, kde jsme označili $\frac{b}{a} = \varphi$. 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.

Kdy skončit algoritmus zlatého řezu? Lze ukázat, že kvůli zaokrouhlovacím chybám nejde interval neurčitosti zmenšit na méně než asi $\sqrt{\varepsilon}$, kde ε je strojová přesnost.

Metoda půlení intervalu a metoda zlatého řezu jsou určeny pouze pro funkce jedné proměnné. Obě konvergují lineárně. To se dá ukázat z toho, že se v každé iteraci interval neurčitosti zmenší o jistý konstantní poměr, v prvním případě $\frac{1}{2}$, ve druhém φ.

9.3 Sestupné metody

Iterační algoritmy na hledání lokálního minima funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ mají tvar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \, \mathbf{s}_k, \tag{9.2}$$

kde vektor $\mathbf{s}_k \in \mathbb{R}^n$ je **směr hledání** a skalár $\alpha_k > 0$ je **délka kroku**. Ve velké třídě algoritmů zvaných **sestupné metody** (*descent methods*) hodnota funkce monotonně klesá, tedy $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$. Existují ale 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*).

Směr $\mathbf{s}_k$ se nazývá **sestupný**, jestliže

$$f'(\mathbf{x}_k)\, \mathbf{s}_k < 0, \tag{9.3}$$

tedy směrová derivace ve směru $\mathbf{s}_k$ 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?). Našli jsme tedy stacionární bod. V tom případě $\mathbf{x}_k$ může (a skoro vždy je) ale také nemusí být lokální minimum.

Máme-li sestupný směr, optimální délku kroku α_k najdeme minimalizací funkce f na polopřímce z bodu $\mathbf{x}_k$ ve směru $\mathbf{s}_k$. Tedy minimalizujeme funkci jedné proměnné

$$\varphi(\alpha_k) = f(\mathbf{x}_k + \alpha_k \mathbf{s}_k) \tag{9.4}$$

přes všechny $\alpha_k \geq 0$. Tato úloha je v kontextu vícerozměrné optimalizace nazývána *line search*. Úlohu stačí často řešit přibližně. Takovou přibližnou metodu není obtížné vymyslet a proto se jí dále nebudeme zabývat.

Dále uvedeme nějznámější zástupce sestupných metod.

9.4 Gradientní metoda

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

$$\mathbf{s}_k = -f'(\mathbf{x}_k)^T = -\nabla f(\mathbf{x}_k). \tag{9.5}$$

Tento směr je sestupný, což je okamžitě vidět dosazením do (9.3).

Rychlost konvergence gradientní metody je lineární. Konvergence je často pomalá kvůli ‘cik-cak’ chování.

9.4.1 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

$$\min_{\mathbf{x}} f(\mathbf{x}) = \min_{\tilde{\mathbf{x}}} \tilde{f}(\tilde{\mathbf{x}}), \qquad \text{kde} \quad \tilde{f}(\tilde{\mathbf{x}}) = \tilde{f}(\mathbf{A}\mathbf{x}) = f(\mathbf{x}) = f(\mathbf{A}^{-1}\tilde{\mathbf{x}}).$$

Iterace gradientní metody v nových proměnných je

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

Zkoumejme, jaké iteraci to odpovídá v původních proměnných. K tomu potřebujeme vyjádřit (9.6) v proměnných $\mathbf{x}$. Použitím řetězového pravidla odvodíme

$$\tilde{f}'(\tilde{\mathbf{x}}) = \frac{\mathrm{d}\tilde{f}(\tilde{\mathbf{x}})}{\mathrm{d}\tilde{\mathbf{x}}} = \frac{\mathrm{d}\tilde{f}(\tilde{\mathbf{x}})}{\mathrm{d}\mathbf{x}}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tilde{\mathbf{x}}} = \frac{\mathrm{d}f(\mathbf{x})}{\mathrm{d}\mathbf{x}}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tilde{\mathbf{x}}} = f'(\mathbf{x})\mathbf{A}^{-1}.$$

Dosazením za $\tilde{\mathbf{x}}$ a $\tilde{f}'(\tilde{\mathbf{x}})$ do (9.6) a úpravou dostaneme

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

To lze napsat ve tvaru (9.2) se směrem hledání

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

Tento směr se liší od původního směru (9.5) 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 (9.8) je také sestupný. Dosazením (9.5) do (9.3) 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 5.18.

Na vzorec (9.8) se lze dívat ještě obecněji. Je jasné, že směr $\mathbf{s}_k = -\mathbf{C}_k^{-1} f'(\mathbf{x}_k)^T$ 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.

9.5 Newtonova metoda

Newtonova metoda (také zvaná Newtonova-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 teď popíšeme.

9.5.1 Použití na soustavy nelineárních rovnic

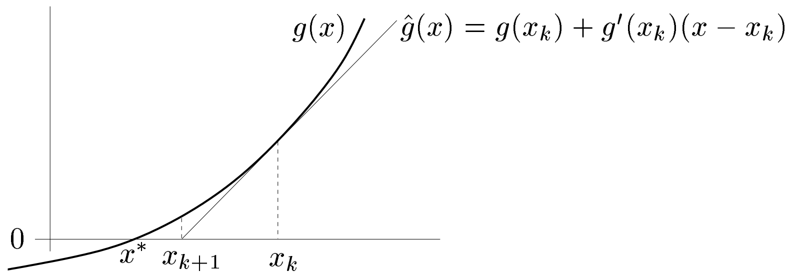
Řešme rovnici $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^n$ je diferencovatelné zobrazení. Jedná se tedy o soustavu n rovnic s n neznámými, které obecně mohou být nelineární. Zobrazení $\mathbf{g}$ aproximujeme v okolí bodu $\mathbf{x}_k$ Taylorovým polynomem prvního řádu

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

kde $\mathbf{g}'(\mathbf{x}_k)$ je derivace zobrazení v bodě $\mathbf{x}_k$, tedy (Jacobiho) matice rozměru $n \times n$. 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

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

Viz obrázek:



Příklad 9.2. *Babylónská metoda* na výpočet druhé odmocniny čísla $a \geq 0$ je dána iterací

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

To není nic jiného než Newtonova metoda pro řešení rovnice $0 = g(x) = x^2 - a$. Opravdu,

$$x_{k+1} = x_k - \frac{g(x)}{g'(x)} = x_k - \frac{x^2 - a}{2x} = x_k - \frac{1}{2}\Big(x_k - \frac{a}{x_k}\Big) = \frac{1}{2}\Big(x_k + \frac{a}{x_k}\Big).$$ □

Příklad 9.3. 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 (9.10) je

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

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 má již 15 platných cifer, $(x_6, y_6) = (\,0.671859751039018,\ 0.944629015546222\,)$. □

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 nebo konverguje k něčemu jinému než k řešení soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$.

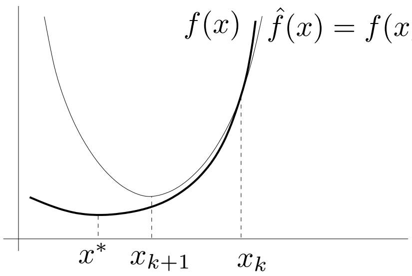
9.5.2 Použití na minimalizaci funkce

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

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

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

Význam iterace (9.10) byl takový, že se zobrazení $\mathbf{g}$ aproximovalo Taylorovým polynomem prvního řádu (tedy afinním zobrazením) a pak se našel kořen tohoto polynomu. Význam iterace (9.11) je takový, že se funkce f aproximuje Taylorovým polynomem druhého řádu (tedy kvadratickou funkcí) a pak se najde minimum této kvadratické funkce. Odvoďte podrobně, že tomu tak je!



Iteraci (9.11) lze napsat v obecnějším tvaru (9.2), kde

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

Výhodou tohoto zobecnění je možnost zvolit optimální (ne nutně jednotkovou) délku kroku pomocí jednorozměrné minimalizace (9.4). Algoritmu (9.11) s jednotkovou délkou kroku se pak říká **čistá** Newtonova metoda.

Vektoru (9.12) říkáme **Newtonův směr**. Aby to byl sestupný směr, musí být

$$f'(\mathbf{x}_k)\,\mathbf{s}_k = -f'(\mathbf{x}_k)\,f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^T < 0.$$

Postačující podmínkou pro to je, aby Hessova matice $f''(\mathbf{x}_k)$ (a tedy i její inverze) byla pozitivně definitní.

9.6 Gaussova-Newtonova metoda

Řešme soustavu rovnic $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ pro $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$, tedy soustavu m rovnic s n neznámými. Dovolíme $m \geq n$, tj. soustava může být přeurčená. Přibližné řešení ve smyslu nejmenších čtverců vede na minimalizaci funkce

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

kde g_i jsou složky zobrazení $\mathbf{g}$. Speciálním případem je přibližné řešení přeurčené lineární nehomogenní soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}$, čemuž jsme se věnovali v §4. Zde ovšem předpokládáme obecně nelineární zobrazení $\mathbf{g}$.

Všimněte si, že zatímco v §9.4 a §9.5.2 bylo cílem minimalizovat *obecnou* funkci, zde chceme minimalizovat funkci ve speciálním tvaru (9.13). Nyní máme dvě možnosti. Buď můžeme nasadit na funkci (9.13) jednu z metod pro minimalizaci obecné funkce, k čemuž se vrátíme v §9.6.1. Nebo můžeme být chytřejší a využít speciálního tvaru funkce (9.13), což uděláme teď.

Aproximujme opět zobrazení $\mathbf{g}$ Taylorovým polynomem prvního řádu (9.9). Úloha (9.13) pak vyžaduje minimalizovat $\|\hat{\mathbf{g}}(\mathbf{x})\|^2$. To je úloha lineárních nejmenších čtverců, kterou jsme vyřešili v §4. Vede na normální rovnice

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

Pokud má Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ plnou hodnost, řešíme pomocí pseudoinverze:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \underbrace{[\mathbf{g}'(\mathbf{x}_k)^T\,\mathbf{g}'(\mathbf{x}_k)]^{-1}\,\mathbf{g}'(\mathbf{x}_k)^T}_{\mathbf{g}'(\mathbf{x}_k)^+}\,\mathbf{g}(\mathbf{x}_k) \tag{9.14}$$

Algoritmus (9.14) je znám jako **Gaussova-Newtonova metoda**. Můžeme jej opět napsat obecněji ve tvaru (9.2) se směrem hledání

$$\mathbf{s}_k = -[\mathbf{g}'(\mathbf{x}_k)^T\,\mathbf{g}'(\mathbf{x}_k)]^{-1}\,\mathbf{g}'(\mathbf{x}_k)^T\,\mathbf{g}(\mathbf{x}_k). \tag{9.15}$$

Pro $m = n$ máme $\mathbf{g}'(\mathbf{x}_k)^+ = \mathbf{g}'(\mathbf{x}_k)^{-1}$, tedy Gaussova-Newtonova metoda se redukuje na Newtonovu metodu (9.10) na řešení soustavy n rovnic s n neznámými.

Snadno spočítáme (viz 7.6) derivaci účelové funkce (9.13), je rovna $f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$. Z toho vidíme, že Gaussův-Newtonův směr (9.15) lze psát ekvivalentně jako

$$\mathbf{s}_k = -\tfrac{1}{2}[\mathbf{g}'(\mathbf{x}_k)^T\,\mathbf{g}'(\mathbf{x}_k)]^{-1}\,f'(\mathbf{x}_k)^T. \tag{9.16}$$

Tento směr se liší od gradientního směru (9.5) 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 hodnost (tedy n), tato matice je pozitivně definitní. Podobnou úvahou jako v §9.4.1 dostaneme, poněkud překvapivě, že směr (9.15) je *vždy* sestupný.

‘Čistá’ Gaussova-Newtonova metoda (tj. s jednotkovou délkou kroku) může divergovat. Protože ale Gaussův-Newtonův směr je vždy sestupný, alespoň přibližnou optimalizací délky kroku α_k lze zajistit konvergenci.

9.6.1 Rozdíl proti plné Newtonově metodě

Předpokládejme, že bychom optimalizovali naší účelovou funkci (9.13) přímo Newtonovou metodou z §9.5.2. Spočítejme (proveďte i sami!) Hessián funkce (9.13):

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

Hessián je součtem členu obsahujícího derivace prvního rádu a členu obsahujícího derivace druhého řádu. Vidíme, že směr (9.16) se liší od Newtonova směru (9.12) zanedbáním členu druhého řádu v Hessiánu $f''(\mathbf{x}_k)$. To se projevuje tím, že Gaussova-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.

9.6.2 Levenbergova-Marquardtova metoda

Levenbergova-Marquardtova metoda je široce používané vylepšení Gaussovy-Newtonovy metody, které matici $\mathbf{g}'(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$ v iteraci (9.14) nahrazuje maticí

$$\mathbf{g}'(\mathbf{x}_k)^T\mathbf{g}'(\mathbf{x}_k) + \mu_k\mathbf{I} \tag{9.17}$$

pro $\mu_k > 0$. Vidíme, že:

- Pro malé μ_k se Levenbergova-Marquardtova iterace blíží Gaussově-Newtonově iteraci.
- Pro velké μ_k je inverze matice (9.17) blízká $\mu_k^{-1}\mathbf{I}$, tedy Levenbergova-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 Gaussovy-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, která v algoritmu nevystupuje.

Na algoritmus lze pohlížet i jinak. V iteraci (9.14) se počítá inverze matice $\mathbf{g}'(\mathbf{x}_k)^T\mathbf{g}'(\mathbf{x}_k)$. Tato matice je sice vždy positivně semidefinitní, ale může být blízká singulární (kdy se to stane?). To neblaze ovlivní stabilitu algoritmu. Matice (9.17) je ale vždy pozitivně definitní (viz Cvičení 5.19), a tedy regulární.

Poznámka na závěr. Probrané algoritmy jsou vlastně velmi jednoduché, je to zcela minimální základ numerické optimalizace. Hrozí však nebezpečí, že se v látce student 'zamotá'. Musí vám být zcela jasné, jakou úlohu každý algoritmus řeší! Udělejte si tabulku, ve které pro každý algoritmus napíšete jednak účelovou funkci a jednak iteraci ve všech možných tvarech. Přemýšlejte o tabulce a hledejte mezi algoritmy souvislosti.

9.7 Cvičení

9.1. Najděte lokální extrém funkce $f(x, y) = x^2 - y + \sin(y^2 - 2x)$ čistou Newtonovou metodou. Počáteční odhad zvolte $(x_0, y_0) = (1, 1)$.

9.2. V systému GPS máme m satelitů se známými souřadnicemi $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ a chceme spočítat naše souřadnice $\mathbf{x} \in \mathbb{R}^n$ z naměřených vzdáleností $y_i = \|\mathbf{a}_i - \mathbf{x}\|$. Tato měření jsou ovšem zatížena chybou, proto obecně nebude žádné $\mathbf{x}$ vyhovující těmto rovnicím. Řešme tuto přeurčenou nelineární soustavu ve smyslu nejmenších čtverců, tedy minimalizujme funkci

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

Odvoďte iteraci Gaussovy-Newtonovy a Levenbergovy-Marquardtovy metody.

9.3. Máme m bodů v rovině o souřadnicích (x_i, y_i), $i = 1, \ldots, m$. 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 Gaussovy-Newtonovy a Levenbergovy-Marquardtovy metody.

9.4. Mějme křivky $(x-3)^2 + y^2 = 1$ a $x^4 + y^4 = 1$. Tyto křivky nemají společný průsečík, tedy soustava rovnic je přeurčená, i když má stejný počet rovnic jako neznámých. Řešte ji Gauss-Newtonovou metodou. Křivky si načrtněte a podle toho zvolte počáteční odhad. Odpovězte na následující otázku: pokud metoda konverguje ke globálnímu maximu, bude toto maximum bod, který minimalizuje součet čtverců vzdáleností k oběma křivkám?

Kapitola 10

Konvexní množiny

10.1 Čtyři kombinace a čtyři obaly

Nechť V je lineární prostor. Vážený součet $\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k$ vektorů $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$ se nazývá jejich

lineární kombinace,	jestliže	$\alpha_1, \ldots, \alpha_k \in \mathbb{R}$.		
afinní kombinace,	jestliže	$\alpha_1, \ldots, \alpha_k \in \mathbb{R}$,	$\alpha_1 + \cdots + \alpha_k = 1$.	
nezáporná kombinace,	jestliže	$\alpha_1, \ldots, \alpha_k \in \mathbb{R}$,		$\alpha_1, \ldots, \alpha_k \geq 0$.
konvexní kombinace,	jestliže	$\alpha_1, \ldots, \alpha_k \in \mathbb{R}$,	$\alpha_1 + \cdots + \alpha_k = 1$,	$\alpha_1, \ldots, \alpha_k \geq 0$.

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**.

Lineární [afinní, nezáporný, konvexní] **obal** vektorů $\mathbf{x}_1, \ldots, \mathbf{x}_k$ je množina všech jejich lineárních [afinních, nezáporných, konvexních] kombinací. Obecněji, obal (konečné či nekonečné) množiny $X \subseteq V$ je množina kombinací všech konečných podmnožin X. Ekvivalentní definice obalů: Lineární [afinní, nezáporný, konvexní] obal množiny $X \subseteq \mathbb{R}^n$ je nejmenší lineární podprostor [afinní podprostor, konvexní kužel, konvexní množina] obsahující množinu X.

Cvičení: Nakreslete lineární, afinní, nezáporný a konvexní obal k bodů v $\mathbb{R}^n$ pro devět případů $k, n \in \{1, 2, 3\}$.

10.2 Konvexní množiny

Zdaleka nejdůležitější nový pojem je konvexní množina. Věnujme se jí podrobněji.

Množina $U \subseteq V$ se nazývá **konvexní množina**, pokud je uzavřená vůči konvexním kombinacím, tj. každá konvexní kombinace bodů z U patří do U. Toto je (což lze dokázat indukcí) ekvivalentní tvrzení, že každá konvexní kombinace *dvojice* bodů z U patří do U. Jinými slovy, množina je konvexní, pokud každá úsečka s krajními body v množině leží celá v množině.

Věta 10.1. *Průnik (i nekonečně mnoha) konvexních množin je konvexní množina.*

Důkaz je snadný.

Sjednocení konvexních množin ale *nemusí* být konvexní množina.

10.3 Konvexní polyedry

(Uzavřený) **poloprostor** je množina $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} \geq b \,\}$ pro nějaké $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$.

Definice 10.1. **Konvexní polyedr** *je průnik konečně mnoha poloprostorů, tedy množina*

$$X = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b} \,\}.$$

Tato definice samozřejmě dovoluje i omezení typu rovnosti $\mathbf{a}^T\mathbf{x} = b$, které jdou složit ze dvou omezení $\mathbf{a}^T\mathbf{x} \leq b$ a $\mathbf{a}^T\mathbf{x} \geq b$. Pak je polyedr průnikem poloprostorů a nadrovin. Nadrovina je ale také průnik dvou poloprostorů.

Jelikož poloprostor je očividně konvexní množina, plyne konvexita konvexního polyedru z Věty 10.1. Všimněte si, že konvexní polyedr nemusí být omezený.

Příklad 10.1. Příklady konvexních polyedrů:

- Prázdná množina $\emptyset$.
- Celý prostor $\mathbb{R}^n$.
- Jediný bod.
- Přímka, ať už procházející nebo neprocházející počátkem.
- Nadrovina $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} = b \,\}$.
- Afinní podprostor.
- Polopřímka $\{\, \mathbf{x}_0 + \alpha\mathbf{s} \mid \alpha \geq 0 \,\}$.
- Poloprostor. Buď uzavřený $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} \geq b \,\}$ nebo otevřený $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} > b \,\}$.
- Pás $\{\, \mathbf{x} \in \mathbb{R}^n \mid b_1 \leq \mathbf{a}^T\mathbf{x} \leq b_2 \,\}$.
- Platónova tělesa v $\mathbb{R}^3$.
- Hyperkrychle $\{\, \mathbf{x} \in \mathbb{R}^n \mid -1 \leq x_i \leq 1,\ i = 1, \ldots, n \,\}$.
- 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 $\{\, \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n |x_i| \leq 1 \,\}$. □

10.3.1 Dvě reprezentace polyedru

Bod $\mathbf{x} \in X$ nazveme *extrémní bod* konvexní množiny X, pokud není roven konvexní kombinaci jiných bodů množiny X. Extrémní body konvexního poleydru nazýváme jeho **vrcholy**.

Věta 10.2. *Konvexní obal konečně mnoha bodů je omezený konvexní polyedr. Obráceně, omezený*[1] *konvexní polyedr je konvexním obalem svých vrcholů.*

Máme tedy dvě reprezentace omezeného polyedru:

[1] Pro neomezené konvexní polyedry platí podobná věta, kterou zde nebudeme uvádět. V ní kromě vrcholů vystupují *paprsky*, což jsou polopřímky, podél nichž polyedr ubíhá do nekonečna.

- **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 velmi těžký nebo i nemožný. Důvodem je to, že polyedr definovaný jako průnik malého počtu (přesněji, tento počet je polynomiální funkcí dimenze n) poloprostorů může mít obrovský (exponenciální v dimenzi n) počet vrcholů. Naopak, polyedr s malým počtem vrcholů může mít exponenciální počet facet. Tedy algoritmus, který převádí H-reprezentaci na V-reprezentaci nebo naopak, by při polynomiálně dlouhém vstupu musel vydat exponenciálně dlouhý výstup.

Příklad 10.2.

- Simplex (tedy konvexní obal $n+1$ bodů) je konvexní polyedr, který má $n+1$ vrcholů a $n+1$ facet.
- Hyperkrychle má $2n$ facet, ale 2^n vrcholů.
- Zobecněný osmistěn $\{\, \mathbf{x} \in \mathbb{R}^n \mid \sum_i |x_i| \leq 1 \,\}$ má $2n$ vrcholů, ale 2^n facet (je to v jistém smyslu duální polyedr k hyperkrychli). □

10.4 Cvičení

10.1. 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)?

a) $\{\, 2y_1 + 3y_2 \mid -1 \leq y_1 \leq 1,\ -1 \leq y_2 \leq 1 \,\}$
b) $\left\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0},\ \mathbf{1}^T\mathbf{x} = 1,\ \sum_i x_1 a_i = b_1,\ \sum_i x_i {a_i}^2 = b_2 \,\right\}$, kde a_i, b_1, b_2 jsou dané skaláry
c) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\| \leq \|\mathbf{x} - \mathbf{b}\| \,\}$, kde $\mathbf{a}, \mathbf{b}$ jsou dány
d) $\{\, \mathbf{Cy} \mid \mathbf{y} \geq \mathbf{0} \,\}$, kde matice $\mathbf{C}$ je dána
e) $\{\, \mathbf{Cy} \mid \|\mathbf{y}\| \leq 1 \,\}$, kde matice $\mathbf{C}$ je dána

10.2. Dokažte z definice konvexní množiny, že následucící množiny jsou konvexní:

a) interval $\langle a, b \rangle \subseteq \mathbb{R}$
b) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \leq \mathbf{b},\ \mathbf{Cx} = \mathbf{d} \,\}$
c) Množina všech positivně definitních matic rozměru $n \times n$.
d) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T\mathbf{Ax} \leq 1 \,\}$, kde $\mathbf{A}$ je pozitivně semidefinitní

10.3. Mějme m vektorů $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$. Pro každé $i = 1, \ldots, m$ definujme množinu

$$X_i = \{\, \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}_i\| \leq \|\mathbf{x} - \mathbf{a}_j\|\ \forall j \neq i \,\}.$$

Dokažte, že množiny $X_1, \ldots, X_m$ jsou konvexní. Ukažte, že tyto množiny tvoří rozklad množiny $\mathbb{R}^n$.

10.4. Které z následujících množin jsou konvexní? Pokud to jde, zkuste množinu načrtnout v prostoru malé dimenze.

a) $\{\, \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1 \,\}$
(Řešení: Nadrovina, tedy konvexní.)

b) $\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i \geq 1 \}$
(Řešení: Poloprostor, tedy konvexní.)

c) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0},\ \sum_{i=1}^n x_i = 1 \}$
(Řešení: Průnik poloprostorů a nadroviny, tedy konvexní polyedr.)

d) $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0},\ \sum_{i=1}^n x_i \leq 1 \}$
(Řešení: Průnik poloprostorů, tedy konvexní polyedr.)

e) $\mathbb{R}^n \setminus \{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i \geq 1 \}$
(Řešení: Doplněk uzavřeného poloprostoru, tedy otevřený poloprostor, tedy konvexní.)

f) $\mathbb{R}^n \setminus \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geq \mathbf{0},\ \sum_{i=1}^n x_i \leq 1 \}$
(Řešení: Doplněk simplexu, není konvexní.)

g) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1 \}$
(Řešení: Sféra (koule bez vnitřku), není konvexní.)

h) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| < 1 \}$
(Řešení: Otevřená koule, konvexní.)

i) $\{ \mathbf{A} \in \mathbb{R}^{n\times n} \mid \mathbf{A} = -\mathbf{A}^T \}$ kde $\mathbb{R}^{n\times n}$ značí množinu všech matic $n \times n$
(Řešení: Lineární podprostor prostoru $\mathbb{R}^{n\times n}$, konvexní.)

j) $\{ (x, y) \in \mathbb{R}^2 \mid x \geq 0,\ y \geq 0,\ xy = 1 \}$
(Řešení: Graf jedné větve hyperboly, není konvexní množina).

k) $\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 2 \} \cap \{ (x, y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 \leq 2 \}$
(Řešení: Průnik dvou koulí, konvexní.)

Kapitola 11

Lineární programování

Úloha **lineárního programování** (LP, také zvané lineární optimalizace) znamená minimalizaci či maximalizaci lineární funkce za podmínek afinních rovností a nerovností. Neboli v obecné formulaci (1.3) je funkce f lineární a funkce g_i, h_i jsou afinní.

Množina přípustných řešení úlohy LP je konvexní polyedr. Platí následující věta.

Věta 11.1. *Lineární funkce nabývá na konvexní množině svého minima v jednom nebo více extrémních bodech této množiny.*

Tedy pokud má úloha LP optimální řešení, nabývá se vždy alespoň v jednom vrcholu množiny přípustných řešení.

Příklad 11.1. Pokud úloha LP má pouze dvě proměnné, můžeme ji řešit graficky pomocí obrázku v rovině. Zkuste to např. pro úlohu

$$
\begin{array}{rrcrcl}
\min & 3x_1 & + & 4x_2 & & \\
\text{za podmínek} & x_1 & + & 2x_2 & \leq & 14 \\
& 3x_1 & - & x_2 & \geq & 0 \\
& x_1 & - & x_2 & \leq & 2
\end{array}
$$

□

11.1 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 takto:

$$
\begin{array}{rrcl}
\min\ (\text{nebo}\ \max) & c_1x_1 + \cdots + c_nx_n & & \\
\text{za podmínek} & a_{i1}x_1 + \cdots + a_{in}x_n \geq b_i, & & i \in I_+ \\
& a_{i1}x_1 + \cdots + a_{in}x_n \leq b_i, & & i \in I_- \\
& a_{i1}x_1 + \cdots + a_{in}x_n = b_i, & & i \in I_0 \\
& x_j \geq 0\,, & & i \in J_+ \\
& x_j \leq 0\,, & & i \in J_- \\
& x_j \in \mathbb{R}, & & j \in J_0
\end{array}
$$

Zde $I = \{1,\dots,m\} = I_0 \cup I_+ \cup I_-$ a $J = \{1,\dots,n\} = J_0 \cup J_+ \cup J_-$ 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.

$$\begin{aligned} \min \quad & c_1x_1 + \cdots + \ c_nx_n \\ \text{za podmínek} \quad & a_{i1}x_1 + \cdots + a_{in}x_n = b_i, \quad i = 1,\dots,m \\ & x_j \geq 0\,, \quad j = 1,\dots,n \end{aligned}$$

  To lze psát maticově jako $\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$, kde $\mathbf{x}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m\times n}$. Tomuto tvaru se říká *standardní*. Podotkněme ovšem, že názvosloví jednotlivých tvarů LP není jednotné, slova jako 'standardní tvar', 'základní tvar' či 'kanonický tvar' tedy mohou znamenat v každé knize něco jiného.

- Tvar $\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$.
- Tvar $\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b},\ \mathbf{x} \in \mathbb{R}^n \,\}$.

Tyto speciální tvary nemají menší vyjadřovací schopnost než obecný tvar, neboť obecný tvar se dá na libovolný speciální tvar převést pomocí následujících operací:

- Minimalizaci nahradíme maximalizací, neboť $\min_{x\in X} f(\mathbf{x}) = -\max_{x\in X}[-f(\mathbf{x})]$.
- Rovnost $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ nahradíme dvěma nerovnostmi $a_{i1}x_1 + \cdots + a_{in}x_n \leq b_i$ a $a_{i1}x_1 + \cdots + a_{in}x_n \geq b_i$.
- Nerovnost $a_{i1}x_1 + \cdots + a_{in}x_n \leq b_i$ převedeme na rovnost přidáním pomocné **slackové proměnné**[1] $u_i \geq 0$ jako $a_{i1}x_1 + \cdots + a_{in}x_n + u_i = b_i$.

  Jak převedeme nerovnost $a_{i1}x_1 + \cdots + a_{in}x_n \geq b_i$ na rovnost?
- Proměnnou bez omezení $x_i \in \mathbb{R}$ rozdělíme na dvě nezáporné proměnné $x_i^+ \geq 0$, $x_i^- \geq 0$ přidáním podmínky $x_i = x_i^+ - x_i^-$.

11.1.1 Po částech lineární účelová funkce

Mějme účelovou funkci

$$f(\mathbf{x}) = \max_{k=1}^{K} (\mathbf{c}_k^T\mathbf{x} + d_k), \tag{11.1}$$

kde $\mathbf{c}_k \in \mathbb{R}^n$ a $d_k \in \mathbb{R}$. Pro dané $\mathbf{x}$ tedy spočítáme číslo $f(\mathbf{x})$ tak, že vezmeme největší z čísel $\mathbf{c}_k^T\mathbf{x} + d_k$ pro $k = 1,\dots,K$ (viz Cvičení 11.3). Uvažujme úlohu

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{11.2}$$

Toto *není* úloha LP, neboť účelová funkce f není lineární nebo afinní, je pouze po částech afinní (později uvidíme, že je konvexní, neboť maximum afinních funkcí je konvexní funkce). Úloha

[1] *Slack* znamená anglicky např. mezeru mezi zdí a skříní, která není zcela přiražená ke zdi. Termín *slack variable* nemá ustálený český ekvivalent, někdy se překládá jako *skluzová proměnná*.

jde ale převést na LP zavedením pomocné proměnné z:

$$\begin{aligned} \min \quad & z \\ \text{za podmínek} \quad & \mathbf{c}_k^T\mathbf{x} + d_k \leq z, \quad k = 1, \ldots, K \\ & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Zde minimalizujeme přes proměnné $\mathbf{x} \in \mathbb{R}^n$ a $z \in \mathbb{R}$. Rozmyslete si dobře, proč mají obě úlohy stejné optimum!

Příklad 11.2. 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. □

Při těchto převodech je nutná opatrnost: pokud bychom v úloze (11.2) maximalizovali místo minimalizovali, převod na LP by nebyl možný.

Podobně lze často na LP převést úlohy, které obsahují minima a maxima v omezeních.

Příklad 11.3. Úloha vlevo není LP, protože omezení $\max\{x, y\} \leq 1$ není lineární nerovost. Ale dá se převést na LP vpravo:

$$\min\{\, x - y \mid x \geq 0,\ y \geq 0,\ \max\{x, y\} \leq 1 \,\} = \min\{\, x - y \mid x \geq 0,\ y \geq 0,\ x \leq 1,\ y \leq 1 \,\}$$ □

11.2 Některé aplikace LP

Zde uvedeme typické aplikace LP. Zdaleka to ale není výčet všech aplikací, ten je totiž nepřeberný. Dále doporučujeme prostudovat příklady v §3.2 ve skriptech [?].

11.2.1 Optimální výrobní program

Z m druhů surovin vyrábíme n druhů výrobků.

- a_{ij} = množství suroviny druhu i potřebné na výrobu výrobku druhu j
- b_i = množství suroviny druhu i, které máme k dispozici
- c_j = zisk z vyrobení jednoho výrobku druhu j

- x_j = počet vyrobených výrobků druhu j

Úkolem je zjistit, kolik jakých výrobků máme vyrobit, abychom dosáhli největšího zisku. Řešení:

$$\max\left\{ \sum_{j=1}^{n} c_j x_j \;\middle|\; \sum_{j=1}^{n} a_{ij} x_j \le b_i,\ x_j \ge 0 \right\}. \tag{11.3}$$

Příklad 11.4. [?] Pán u stánku prodává lupínky za 120 Kč/kg a hranolky za 76 Kč/kg. Na výrobu 1 kg lupínků se spotřebuje 2 kg brambor a 0.4 kg oleje. Na výrobu 1 kg 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

$$\begin{array}{rrcl} \max & 120\ell + 76h & & \\ \text{za podmínek} & 2\ell + 1.5h & \le & 100 \\ & 0.4\ell + 0.2h & \le & 16 \\ & \ell, h & \ge & 0 \end{array}$$

Přitom předpokládáme, že zbytky surovin se po pracovní době vyhodí. Pokud se zbytky využijí, tak maximalizujeme $(120-24-16)\ell + (76-18-8)h = 80\ell + 50h$. □

11.2.2 Směšovací (dietní) problém

Z n druhů surovin, z nichž každá je směsí m druhů látek, máme namíchat konečný produkt o požadovaném složení tak, aby cena surovin byla minimální.

- a_{ij} = množství látky druhu i obsažené v jednotkovém množství suroviny druhu j
- b_i = požadované množství látky druhu i v konečném produktu
- c_j = jednotková cena suroviny druhu j
- x_j = množství suroviny druhu j

Řešení:

$$\min\left\{ \sum_{j=1}^{n} c_j x_j \;\middle|\; \sum_{j=1}^{n} a_{ij} x_j = b_i,\ x_j \ge 0 \right\}$$

Příklad je uveden ve Cvičení 11.5.

11.2.3 Dopravní problém

Máme m výrobců a n spotřebitelů.

- a_i = množství zboží vyráběné výrobcem i
- b_j = množství zboží požadované spotřebitelem j
- c_{ij} = cena dopravy jednotky zboží od výrobce i ke spotřebiteli j
- x_{ij} = množství zboží vezené od výrobce i ke spotřebiteli j

Chceme co nejlevněji rozvézt zboží od výrobců ke spotřebitelům. Řešení:

$$\min\left\{ \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} x_{ij} \;\middle|\; \sum_{j=1}^{n} x_{ij} = a_i,\ \sum_{i=1}^{m} x_{ij} = b_j,\ x_{ij} \ge 0 \right\}.$$

Zadání musí splňovat $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$ (nabídka musí být rovna poptávce), jinak bude úloha nepřípustná. Úloha jde modifikovat tak, že dovolíme $\sum_{i=1}^m a_i \geq \sum_{j=1}^n b_j$ (proveďte!).

11.2.4 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:

$$\min\left\{ \sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \;\middle|\; \sum_{j=1}^n x_{ij} \leq a_i,\ \sum_{i=1}^m k_{ij}x_{ij} = b_j,\ x_{ij} \geq 0 \right\}.$$

11.3 Řešení přeurčených lineárních soustav

11.3.1 Vektorové normy

Definice 11.1. *Funkce* $\mathbb{R}^n \to \mathbb{R}$ *se nazývá* **norma** *a značí se* $\|\mathbf{x}\|$, *jestliže splňuje tyto axiomy:*

1. *Jestliže* $\|\mathbf{x}\| = 0$ *pak* $\mathbf{x} = \mathbf{0}$.
2. $\|\lambda\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$ *pro každé* $\lambda \in \mathbb{R}$ *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 snadno vyplývají tyto další vlastnosti normy:

- $\|\mathbf{0}\| = 0$, což plyne z homogenity pro $\lambda = 0$
- $\|\mathbf{x}\| \geq 0$ 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 formalizuje pojem 'délky' vektoru $\mathbf{x}$.

Jednotková sféra normy je množina $\{\, \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1 \,\}$ všech vektorů s jednotkovou normou. Díky axiomu homogenity je jednotková sféra středově symetrická a její tvar zcela určuje normu.

Uveďme příklady norem. Základním příkladem je ℓ_p-norma

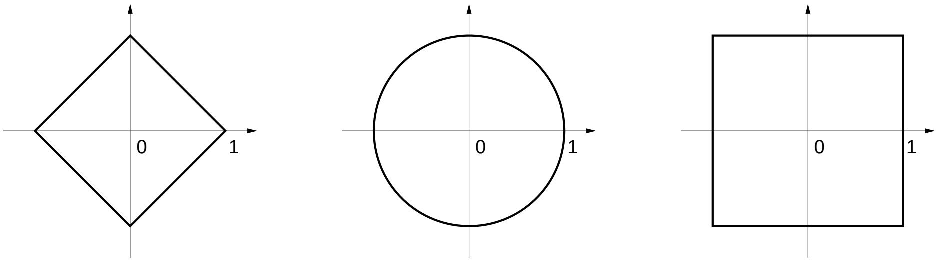
$$\|\mathbf{x}\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}.$$

Musí být $p \geq 1$, jinak neplatí trojúhelníková nerovnost. Nejčastěji narazíte na:

- $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|$. Někdy se jí říká manhattanská norma, protože v systému pravoúhlých ulic je vzdálenost mezi body $\mathbf{x}$ a $\mathbf{y}$ rovna $\|\mathbf{x} - \mathbf{y}\|_1$.
- $\|\mathbf{x}\|_2 = \sqrt{{x_1}^2 + \cdots + {x_n}^2} = \sqrt{\mathbf{x}^T\mathbf{x}}$. Je to známá eukleidovská norma.

- $\|\mathbf{x}\|_\infty = \max\{|x_1|,\ldots,|x_n|\}$ (spočítejte podrobně limitu $\lim_{p\to\infty}\|\mathbf{x}\|_p$!). Nazývá se Chebyševova nebo max-norma.

Jednotkové sféry těchto norem v $\mathbb{R}^2$ vypadají takto:



$\|\mathbf{x}\|_1 = 1$ $\qquad$ $\|\mathbf{x}\|_2 = 1$ $\qquad$ $\|\mathbf{x}\|_\infty = 1$

Existují ale i normy, které nejsou ℓ_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 hodností, je také $\|\mathbf{Ax}\|$ norma.

11.3.2 Řešení přeurčených soustav

Nechť $\mathbf{A} \in \mathbb{R}^{m\times n}$, kde $m > n$. Řešme přeurčenou nehomogenní lineární soustavu $\mathbf{Ax} = \mathbf{b}$ jako

$$\min\{\, \|\mathbf{Ax} - \mathbf{b}\|_p \mid \mathbf{x} \in \mathbb{R}^n \,\} \tag{11.4}$$

(definice ℓ_p normy je v §11.3.1). Pro $p = 2$ dostaneme řešení ve smyslu nejmenších čtverců, kterým jsme se již zabývali v §4.

Pro $p = \infty$ hledáme takové $\mathbf{x}$, které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_\infty = \max_{i=1}^{m} |\mathbf{a}_i^T\mathbf{x} - b_i|, \tag{11.5}$$

tedy minimalizuje maximální residuum. Toto řešení je známé pod názvem *minimaxní* nebo *Čebyševovo*. Tato úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & \mathbf{a}_i^T\mathbf{x} - b_i \le z, \quad i = 1,\ldots,m \\ & -\mathbf{a}_i^T\mathbf{x} + b_i \le z, \quad i = 1,\ldots,m \end{aligned}$$

který lze zapsat elegantněji jako

$$\min\{\, z \in \mathbb{R} \mid \mathbf{x} \in \mathbb{R}^n,\ -z\mathbf{1} \le \mathbf{Ax} - \mathbf{b} \le z\mathbf{1} \,\}. \tag{11.6}$$

Pro $p = 1$ hledáme takové $\mathbf{x}$, které minimalizuje výraz

$$\|\mathbf{Ax} - \mathbf{b}\|_1 = \sum_{i=1}^{m} |\mathbf{a}_i^T\mathbf{x} - b_i|, \tag{11.7}$$

Úloha je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} z_i \\ \text{za podm.} \quad & \mathbf{a}_i^T\mathbf{x} - b_i \le z_i, \quad i = 1,\ldots,m \\ & -\mathbf{a}_i^T\mathbf{x} + b_i \le z_i, \quad i = 1,\ldots,m \end{aligned}$$

který lze zapsat elegantněji jako

$$\min\{\, \mathbf{1}^T\mathbf{z} \mid \mathbf{z} \in \mathbb{R}^m,\ \mathbf{x} \in \mathbb{R}^n,\ -\mathbf{z} \le \mathbf{Ax} - \mathbf{b} \le \mathbf{z} \,\}. \tag{11.8}$$

11.3.3 Použití na robustní regresi

Řešení ve smyslu ℓ_1 normy se používá tehdy, když potřebujeme modelovat funkční závislost naměřených dat (tedy děláme regresi, viz §4.2) 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 ℓ_1 normy tuto neblahou vlastnost nemá, přesněji, má ji v menší míře.

Ukážeme to na nejjednodušším možném případu regrese: odhad hodnoty jediného reálného čísla ze souboru jeho nepřesných měření. Mějme čísla $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$ a řešme úlohu (11.4) ve tvaru

$$\min\{\, \|\mathbf{1}x - \mathbf{y}\|_p \mid x \in \mathbb{R}\,\}. \tag{11.9}$$

- Pro $p = \infty$ minimalizujeme funkci $f(x) = \max_{i=1}^m |x - y_i|$. Řešením střed intervalu krajních bodů, $x = \frac{1}{2}\bigl(\min_{i=1}^m y_i + \max_{i=1}^m y_i\bigr)$.
- Pro $p = 2$ minimalizujeme funkci $f(x) = \sqrt{\sum_{i=1}^m (x - y_i)^2}$. Řešením je aritmetický průměr, $x = \frac{1}{m}\sum_{i=1}^m y_i$ (viz Příklad 4.2).
- Pro $p = 1$ minimalizujeme funkci $f(x) = \sum_{i=1}^m |x - y_i|$. Řešením x je *medián* z čísel y_i. Medián se vypočte tak, že seřadíme čísla y_i podle velikosti a vezmeme prostřední z nich. Pokud je m sudé, máme dva prostřední prvky a v tom případě funkce f nabývá minima v jejich libovolné konvexní kombinaci. Je pak úzus definovat medián jako aritmetický průměr prostředních prvků.

Předpokládejme nyní, že jeden libovolný bod (např. y_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 y_1 dosáhneme *libovolné* hodnoty x. Pro medián to ovšem neplatí – zvětšováním jediného bodu y_i ovlivníme x jen natolik, nakolik to změní pořadí bodů. Jeho libovolným zvětšováním nedosáhneme libovolné hodnoty x.

Příklad 11.5. Šuplérou změříme průměr ocelové kuličky v několika místech, dostaneme hodnoty $\mathbf{y} = (1.02, 1.04, 0.99, 2.03)$ (cm). Při posledním měření jsme se na stupnici přehlédli, proto je poslední hodnota úplně špatně. Z těchto měření chceme odhadnout skutečný průměr. Máme

$$\operatorname*{median}_{i=1}^{m} y_i = 1.005, \qquad \frac{1}{m}\sum_{i=1}^{m} y_i = 1.27, \qquad \frac{1}{2}\bigl(\min_{i=1}^{m} y_i + \max_{i=1}^{m} y_i\bigr) = 1.51.$$

Je zjevné, že medián je neovlivněn vychýleným bodem, zatímco ostatní odhady ano. □

Ve složitějším případě, např. prokládání dat polynomem jako v Příkladu 4.2, se nedá robustnost ℓ_1 řešení takto jednoduše formálně ukázat. Ale intuitivně bude situace obdobná: ℓ_1 řešení bude méně citlivé na vychýlené body než ℓ_2 řešení.

11.4 Cvičení

11.1. Najděte graficky množinu optimálních řešení úlohy

$$\begin{aligned} \min \quad & c_1x_1 + c_2x_2 + c_3x_3 \\ \text{za podm.} \quad & x_1 + x_2 \geq 1 \\ & x_1 + 2x_2 \leq 3 \\ & x_1 + x_2 \leq 10 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

pro následující případy: $\mathbf{c} = (-1, 0, 1)$, $\mathbf{c} = (0, 1, 0)$, $\mathbf{c} = (0, 0, -1)$.

11.2. Vyřešte úvahou tyto úlohy a napište vzorec pro optimální hodnotu. Ve všech úlohách optimalizujeme přes proměnné $\mathbf{x} \in \mathbb{R}^n$ (příp. také $\mathbf{y} \in \mathbb{R}^n$). Parametry $\mathbf{c} \in \mathbb{R}^n$ a $k \in \mathbb{N}$, $1 \leq k \leq n$, jsou dány.

a) $\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \,\}$
(výsledek: $\sum_{i \mid c_i < 0} c_i$, tedy součet těch čísel c_i která jsou záporná)

b) $\min\{\, \mathbf{c}^T\mathbf{x} \mid -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1} \,\}$
(výsledek: $\sum_i |c_i|$)

c) ($\star$) $\min\{\, \mathbf{c}^T\mathbf{x} \mid 0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq 1 \,\}$
(nápověda: proveďte substituci proměnných $y_i = x_i - x_{i-1}$)

d) $\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \geq \mathbf{0},\ \mathbf{1}^T\mathbf{x} = 1 \,\}$

e) $\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \geq \mathbf{0},\ \mathbf{1}^T\mathbf{x} \leq 1 \,\}$

f) $\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \geq \mathbf{0},\ \mathbf{1}^T\mathbf{x} = k \,\}$

g) $\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \geq \mathbf{0},\ \mathbf{1}^T\mathbf{x} \leq k \,\}$

h) ($\star$) $\max\{\, \mathbf{c}^T\mathbf{x} \mid -\mathbf{y} \leq \mathbf{x} \leq \mathbf{y},\ \mathbf{1}^T\mathbf{y} = k,\ \mathbf{y} \leq \mathbf{1} \,\}$

i) ($\star$) $\min\{\, \mathbf{1}^T(\mathbf{u} + \mathbf{v}) \mid \mathbf{u}, \mathbf{v} \geq \mathbf{0},\ \mathbf{u} - \mathbf{v} = \mathbf{c} \,\}$

j) ($\star$) $\min\{\, \mathbf{a}^T\mathbf{u} + \mathbf{b}^T\mathbf{v} \mid \mathbf{u}, \mathbf{v} \geq \mathbf{0},\ \mathbf{u} - \mathbf{v} = \mathbf{c} \,\}$, kde $\mathbf{a}, \mathbf{b}$ jsou dány a platí $\mathbf{a} \geq \mathbf{b}$

11.3. Pochopte následující kód, který vizualizuje funkci $f(\mathbf{x}) = \max_{k=1}^{K}(\mathbf{c}_k^T\mathbf{x} + d_k)$ 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
```

11.4. Převeďte na LP nebo odůvodněte, proč to nejde:

a) $\max\{\, |\mathbf{c}^T\mathbf{x}| \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$

b) $\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ |\mathbf{d}^T\mathbf{x} + e| \leq f,\ \mathbf{x} \geq \mathbf{0} \,\}$

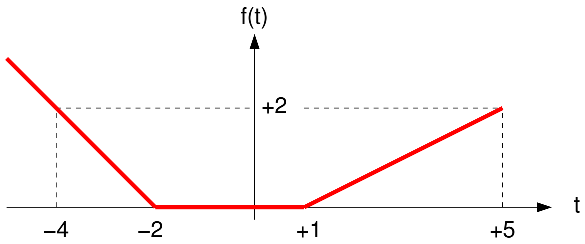
c) $\min\left\{\, \sum_{\ell=1}^{L} \max_{k=1}^{K}(\mathbf{c}_{k\ell}^T\mathbf{x} + d_{k\ell}) \;\middle|\; \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\right\}$

d) $\max\{\, |x - 1| + 2|y + 1| \mid x, y \in \mathbb{R},\ x + y \leq 2 \,\}$

e) $\min\{\, |x_1| + |x_2| + |x_3| \mid 2x_1 - x_2 - x_3 \geq 1,\ -x_1 + 2x_2 - x_3 \geq 1,\ -x_1 - x_2 + 2x_3 \geq 1 \,\}$

f) $\max\{\, \min\{\mathbf{p}^T\mathbf{x},\, \mathbf{q}^T\mathbf{x}\} - |\mathbf{r}^T\mathbf{x}| \;\mid\; \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} \geq \mathbf{b} \,\}$, kde $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \mathbb{R}^n$

g) $\min_{\mathbf{x}\in\mathbb{R}^n} \sum_{i=1}^m f(\mathbf{a}_i^T\mathbf{x} - b_i)$, kde funkce f je dána obrázkem



h) $\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1$ (pro definici p-normy viz 11.3.1)

i) $\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty$

j) $\min\{\, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 \mid \mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{x}\|_\infty \leq 1 \,\}$

k) $\min\{\, \|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty \leq 1 \,\}$

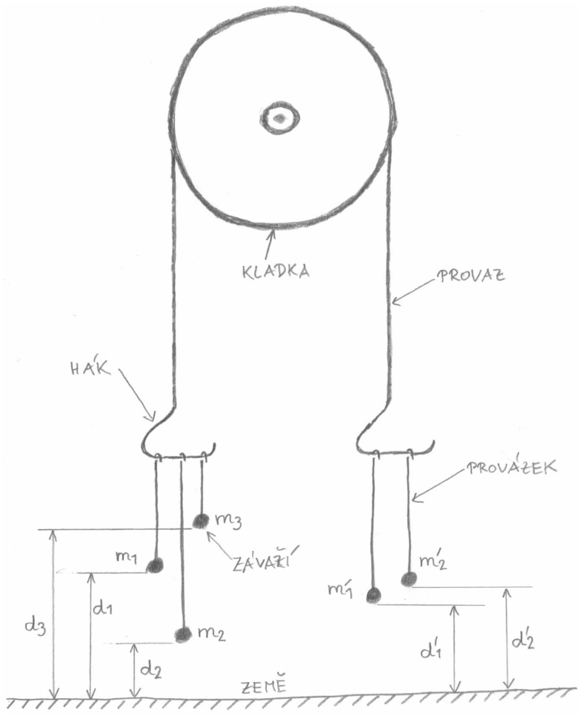
l) $\min_{\mathbf{x}\in\mathbb{R}^n} (\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 + \|\mathbf{x}\|_\infty)$

m) $\min_{\mathbf{x}\in\mathbb{R}^n} \sum_{k=1}^K \|\mathbf{A}_k\mathbf{x} - \mathbf{b}_k\|_\infty$, kde $\mathbf{A}_1, \ldots, \mathbf{A}_K$ jsou dané matice a $\mathbf{b}_1, \ldots, \mathbf{b}_K$ jsou dané vektory.

11.5. Jste kuchařka v menze a máte uvařit pro studenty co nejlevnější oběd, ve kterém ovšem musí být dané minimální množství živin (cukrů, bílkovin a vitamínů). Oběd budete vařit ze tří surovin: brambor, masa a zeleniny. Kolik je třeba každé suroviny na jeden oběd?

	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	

11.6. 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 $i = 1, \ldots, n$. 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 $i = 1, \ldots, n'$. Výšky d_i a d'_i se měří v poloze, kdy jsou oba háky ve stejné výšce nad zemí. Kladka se pohybuje bez tření a provázky mají nulovou hmotnost a stále stejnou délku nezávisle na síle napnutí. Obrázek ukazuje příklad soustavy pro $n = 3$, $n' = 2$.
Je-li to možné, napište lineární program, jehož optimum je rovno minimální potenciální energii soustavy. Není-li to možné, vysvětlete proč.



11.7. Dokažte, že

$$\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{x}\| = 1 \,\} = \max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{x}\| \leq 1 \,\},$$

tj. že omezující podmínka $\|\mathbf{x}\| \le 1$ úlohy vpravo bude v optimu vždy aktivní. Zde $\|\cdot\|$ je libovolná norma a $\mathbf{c} \in \mathbb{R}^n$ je daný vektor.

Nápověda: Jako první krok důkazu si uvědomte, že optimální hodnota bude vždy nezáporná, protože jinak bychom ji mohli zlepšit nahrazením $\mathbf{x}$ za $-\mathbf{x}$.

Platí tvrzení i pro úlohu $\min\{\, \mathbf{c}^T\mathbf{x} \mid \|\mathbf{x}\| = 1 \,\}$? Platí pro úlohu $\max\{\, \|\mathbf{A}\mathbf{x}\| \mid \|\mathbf{x}\| = 1 \,\}$?

Kapitola 12

Simplexová metoda

Zde popíšeme algoritmus na řešení úloh lineárního programování zvaný **simplexová metoda**.

Zapomeňme prozatím na účelovou funkci a zkoumejme množinu přípustných řešení LP ve standardním tvaru

$$X = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}, \tag{12.1}$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$ je široká ($m < n$) matice s hodností m, tedy její řádky jsou lineárně nezávislé.

Soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ má nekonečně mnoho řešení. Položíme-li však $n - m$ složek vektoru $\mathbf{x}$ rovno nule, soustava může mít jediné řešení. Tato úvaha vede k následujícím definicím:

- Množina $L \subseteq \{1, 2, \ldots, n\}$ se nazývá **báze** úlohy, pokud $|L| = m$ a sloupce matice $\mathbf{A}$ s indexy L jsou lineárně nezávislé. Tedy sloupce L tvoří regulární matici $m \times m$.
- Vektor $\mathbf{x}$ se nazývá **bázové řešení** příslušné bázi L, pokud $\mathbf{A}\mathbf{x} = \mathbf{b}$ a $x_j = 0$ pro $j \notin L$.
- Bázové řešení $\mathbf{x}$ se nazývá **přípustné**, pokud $\mathbf{x} \geq \mathbf{0}$.
- Bázové řešení $\mathbf{x}$ se nazývá **degenerované**, pokud má méně než m nenulových složek.
- Dvě báze se nazývají **sousední**, pokud mají $m - 1$ společných prvků.

Protože $\mathbf{A}$ má hodnost m, existuje aspoň jedna báze. Je jasné, že báze určuje jednoznačně bázové řešení. Bázové řešení však může odpovídat více než jedné bázi, což se stane právě tehdy, když je toto bázové řešení degenerované.

Příklad 12.1. Nechť je soustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ dána tabulkou

$$[\,\mathbf{A} \,|\, \mathbf{b}\,] = \left[\begin{array}{rrrrrr|r} -1 & 1 & 3 & 1 & 0 & 2 & 1 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \\ -1 & 0 & 4 & 1 & 1 & 4 & 2 \end{array}\right].$$

- $L = \{2, 3, 5\}$ není báze, protože sloupce $2, 3, 5$ matice $\mathbf{A}$ jsou lineárně závislé.
- $L = \{1, 2, 4\}$ je báze, protože tyto sloupce jsou lineárně nezávislé. Bázové řešení $\mathbf{x} = (x_1, x_2, \ldots, x_6)$ příslušné bázi L se najde řešením soustavy

$$\begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

  a položením $x_3 = x_5 = x_6 = 0$. Dostaneme $\mathbf{x} = (4, -1, 0, 6, 0, 0)$. Toto bázové řešení je nepřípustné, protože $x_2 < 0$. Není degenerované, protože má $m = 3$ nenulových složek.

- $L = \{1,2,6\}$ je báze. Bázové řešení je $\mathbf{x} = (1, \frac{1}{2}, 0, 0, 0, \frac{3}{4})$. Je přípustné, protože $\mathbf{x} \geq \mathbf{0}$.
- $L = \{3,4,5\}$ je báze. Bázové řešení je $\mathbf{x} = (0,0,1,-2,0,0)$. Je nepřípustné. Navíc je degenerované, protože má méně než $m - 1 = 3$ nenulových složek.
- Stejné bázové řešení $\mathbf{x} = (0,0,1,-2,0,0)$ dostaneme volbou báze $L = \{3,4,6\}$. Vidíme, že bázové řešení odpovídá více než jedné bázi, protože je degenerované.
- Báze $\{2,3,5\}$ a $\{3,4,5\}$ jsou sousední, protože mají společné dva prvky $\{3,5\}$. Báze $\{2,3,5\}$ a $\{1,2,4\}$ nejsou sousední, protože mají společný jen jeden prvek. □

Lze ukázat (přesný důkaz vynecháme), že přípustná bázová řešení odpovídají vrcholům konvexního polyedru X. Dále víme, že optimum lineární funkce se musí nabývat alespoň v jednom vrcholu. To nám dovoluje navrhnout naivní algoritmus na řešení LP: uděláme výčet všech přípustných bázových řešení a nalezneme to s nejlepší hodnotou účelové funkce. Tato metoda samozřejmě nelze prakticky použít, protože přípustných bázových řešení je exponenciálně mnoho.

Simplexová metoda je efektivnější obměna tohoto přístupu: přechází mezi sousedními bázemi tak, že bázová řešení jsou stále přípustná a účelová funkce se zlepšuje (nebo aspoň nezhoršuje).

12.1 Stavební kameny algoritmu

Zde vysvětlíme jednotlivé stavební kameny simplexové metody, které nakonec v §12.2 spojíme v celý algoritmus.

12.1.1 Přechod k sousední standardní bázi

Ačkoliv výše uvedená definice báze dovoluje libovolné báze, simplexový algoritmus udržuje pouze *standardní báze.* V tom případě jsou nenulové složky bázového řešení $\mathbf{x}$ rovny jednoduše složkám vektoru $\mathbf{b}$.

Z lineární algebry známe *ekvivalentní řádkové úpravy* soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$: libovolný řádek tabulky $[\,\mathbf{A} \mid \mathbf{b}\,]$ můžeme vynásobit nenulovým číslem a můžeme k němu přičíst libovolnou lineární kombinaci ostatních řádků. Tyto úpravy nemění množinu řešení soustavy.

Ukážeme, jak přejít od aktuální standardní báze L k sousední standardní bázi, tedy nahradit jeden bázový sloupec $j' \in L$ nebázovým sloupcem $j \notin L$. Nechť i je takový řádek, ve kterém má sloupec j' jedničku, $a_{ij'} = 1$. Prvek a_{ij} se nazývá **pivot** (angl. znamená *čep*). Předpokládáme, že $a_{ij} \neq 0$. Chceme nastavit pivot a_{ij} na jedničku, vynulovat prvky nad i pod pivotem, a nezměnit přitom sloupce $L \setminus \{j'\}$. 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 12.2. Mějme soustavu

$$[\,\mathbf{A} \mid \mathbf{b}\,] = \left[\begin{array}{cccccc|c} 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & \boxed{1} & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array}\right]$$

se standardní bází $L = \{1,4,5\}$. Vidíme ihned odpovídající bázové řešení, $\mathbf{x} = (3,0,0,4,1,0)$.

Nahradíme bázový sloupec $j' = 1$ nebázovým sloupcem $j = 2$, tedy přejdeme k sousední bázi $\{2, 4, 5\}$. Máme $i = 2$, tedy pivot je prvek a_{22} (v tabulce zvýrazněn). Řádkovými úpravami musíme docílit, aby pivot byl roven jedné a prvky nad ním a pod ním byly nulové. Při tom smíme 'zničit' sloupec 1, ale sloupce 4 a 5 se změnit nesmějí. Toho se docílí vydělením řádku 2 číslem a_{22} (což zde lze vynechat, protože náhodou máme $a_{22} = 1$) a pak přičtením vhodných násobků řádku 2 k ostatním řádkům. Výsledek:

$$[\,\mathbf{A}\,|\,\mathbf{b}\,] = \left[\begin{array}{rrrrrr|r} -2 & 0 & 0 & 1 & 0 & 0 & -2 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 1 & 0 & 4 & 0 & 1 & 4 & 4 \end{array}\right].$$

Nyní sloupce $\{2, 4, 5\}$ tvoří standardní bázi. □

12.1.2 Kdy je sousední bázové řešení přípustné?

Uvedeným způsobem můžeme od aktuální standardní báze přejít k libovolné sousední standardní bázi. Přitom nové bázové řešení může nebo nemusí být přípustné. Je-li aktuální bázové řešení přípustné, jak poznáme, zda i nové bázové řešení bude přípustné?

Protože nenulové složky bázového řešení $\mathbf{x}$ jsou rovny složkám vektoru $\mathbf{b}$, bázové řešení je přípustné právě tehdy, když $\mathbf{b} \geq \mathbf{0}$. Nechť v aktuální tabulce je $\mathbf{b} \geq \mathbf{0}$. Proveďme 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 §12.1.1):

- b_i se změní na b_i/a_{ij},
- pro $i' \neq i$ se $b_{i'}$ změní na $b_{i'} - a_{i'j}(b_i/a_{ij})$.

Tato čísla musejí být nezáporná. Tvrdíme, že to nastane právě tehdy, když platí následující podmínky:

$$a_{ij} > 0 \tag{12.2a}$$

$$\text{Pro každé } i' \neq i \text{ platí } a_{i'j} \leq 0 \text{ nebo } \frac{b_i}{a_{ij}} \leq \frac{b_{i'}}{a_{i'j}}. \tag{12.2b}$$

('nebo' je zde užito v nevylučovacím smyslu).

Podmínka (12.2a) je zřejmá. Po jistém zamyšlení vidíme, že podmínka (12.2b) je ekvivalentní podmínce $b_{i'} - a_{i'j}(b_i/a_{ij}) \geq 0$, uvědomíme-li si, že $a_{ij} > 0$, $b_i > 0$, $b_{i'} > 0$.

Příklad 12.3. Uvažujme opět naši soustavu

$$[\,\mathbf{A}\,|\,\mathbf{b}\,] = \left[\begin{array}{llllll|l} 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array}\right].$$

- Povede ekvivalentní úprava okolo pivotu $(i, j) = (3, 2)$ k přípustnému řešení? Ne, protože $a_{ij} = -1 < 0$, což porušuje podmínku (12.2a).
- Povede ekvivalentní úprava okolo pivotu $(i, j) = (2, 2)$ k přípustnému řešení (tuto úpravu jsme již provedli v Příkladu 12.2)? Ne, protože (i, j) nesplňuje podmínku (12.2b). Pro $i' = 3$ máme $a_{i'j} = -1 \leq 0$, tedy podmínka (12.2b) je splněna. Ale pro $i' = 1$ je $a_{i'j} > 0$, tedy musí být $\frac{3}{1} \leq \frac{4}{2}$, což není pravda.
- Povede ekvivalentní úprava okolo pivotu $(i, j) = (3, 6)$ k přípustnému řešení? Podmínka (12.2a) je splněna, protože $a_{ij} = 2 > 0$. Podmínka (12.2b) vyžaduje $\frac{1}{2} \leq \frac{4}{4}$ a $\frac{1}{2} \leq \frac{3}{2}$, což platí. □

12.1.3 Co znamená nekladný sloupec?

Jestliže jsou všechny prvky v nebázovém sloupci nekladné, tento sloupec se nemůže stát bázovým, neboť v něm nelze vybrat pivot splňující podmínku (12.2a). V tom případě se některé složky vektoru $\mathbf{x}$ mohou 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 12.4. Nechť tabulka $[\,\mathbf{A}\,|\,\mathbf{b}\,]$ vypadá takto:

$$
\begin{array}{r|rrrrrr|r}
 & 0 & -2 & 6 & 1 & 0 & 4 & 4 \\
 & 1 & -1 & 3 & 0 & 0 & 2 & 3 \\
 & 0 & -1 & 1 & 0 & 1 & 2 & 1 \\
\hline
\mathbf{x} = & 3 & 0 & 0 & 4 & 1 & 0 &
\end{array}
$$

Báze je $\{1, 4, 5\}$. Pod tabulkou je napsáno odpovídající bázové řešení $\mathbf{x}$. Když se x_2 bude libovolně zvětšovat, změnu lze kompenzovat současným zvětšováním x_1, x_4, x_5 tak, že vektor $\mathbf{Ax}$ zůstane nezměněn a tedy roven $\mathbf{b}$. Konkrétně, vektor $\mathbf{x} = (\,3+\alpha,\ \alpha,\ 0,\ 4+2\alpha,\ 1+\alpha,\ 0\,)$ bude pro každé $\alpha \geq 0$ splňovat $\mathbf{Ax} = \mathbf{b}$ a $\mathbf{x} \geq \mathbf{0}$. □

12.1.4 Ekvivalentní úpravy účelového řádku

Dosud jsme prováděli ekvivalentní řádkové úpravy pouze na soustavě $\mathbf{Ax} = \mathbf{b}$ a účelové funkce si nevšímali. 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

$$\min\{\, \mathbf{c}^T\mathbf{x} - d \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\,\}. \tag{12.3}$$

Úlohu budeme reprezentovat **simplexovou tabulkou**

$$\left[\begin{array}{c|c} \mathbf{c}^T & d \\ \hline \mathbf{A} & \mathbf{b} \end{array}\right]. \tag{12.4}$$

Přičtěme k účelovému řádku $[\,\mathbf{c}^T\,|\,d\,]$ libovolnou lineární kombinaci $\mathbf{y}^T[\,\mathbf{A}\,|\,\mathbf{b}\,]$ ostatních řádků $[\,\mathbf{A}\,|\,\mathbf{b}\,]$, kde $\mathbf{y}$ značí koeficienty lineární kombinace. Ukážeme, že tato úprava zachová hodnotu účelové funkce $\mathbf{c}^T\mathbf{x} - d$ pro každé $\mathbf{x}$ splňující $\mathbf{Ax} = \mathbf{0}$. 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

$$(\mathbf{c}^T + \mathbf{y}^T\mathbf{A})\mathbf{x} - (d + \mathbf{y}^T\mathbf{b}) = \mathbf{c}^T - d + \mathbf{y}^T(\mathbf{Ax} - \mathbf{b}).$$

Ale to je rovno $\mathbf{c}^T - d$ pro každé $\mathbf{x}$ splňující $\mathbf{Ax} = \mathbf{b}$.

12.1.5 Co udělá přechod k sousední bázi s účelovou funkcí?

Nechť sloupce L tvoří standardní bázi. Přičtěme k účelovému řádku takovou lineární kombinaci ostatních řádků, aby koeficienty c_j v bázových sloupcích $j \in L$ byly nulové. To lze vždy udělat. Protože bázové řešení $\mathbf{x}$ je v nebázových sloupcích nulové, znamená to $\mathbf{c}^T\mathbf{x} = 0$. Tedy hodnota kritéria $\mathbf{c}^T\mathbf{x} - d$ v bázovém řešení $\mathbf{x}$ je rovna jednoduše $-d$. Navíc je na první pohled vidět, co udělá s kritériem vložení nebázového sloupce $j \notin L$ do báze: při $c_j \geq 0$ kritérium stoupne nebo se nezmění, při $c_j \leq 0$ kritérium klesne nebo se nezmění.

Příklad 12.5. Mějme úlohu se standardní bází $\{1, 4, 5\}$:

$$\left[\begin{array}{c|c} \mathbf{c}^T & d \\ \hline \mathbf{A} & \mathbf{b} \end{array}\right] = \left[\begin{array}{rrrrrr|r} 1 & -2 & -3 & -1 & 2 & 1 & 4 \\ \hline 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ 0 & -1 & 1 & 0 & 1 & 2 & 1 \end{array}\right].$$

Vynulujme hodnoty vektoru $\mathbf{c}$ v bázových sloupcích. To uděláme tak, že ke kriteriálnímu řádku přičteme první řádek, odečteme druhý řádek, a odečteme dvojnásobek třetího řádku:

$$\begin{array}{r|rrrrrr|r} & 0 & 1 & -2 & 0 & 0 & -1 & 3 \\ \hline & 0 & 2 & 6 & 1 & 0 & 4 & 4 \\ & 1 & 1 & 3 & 0 & 0 & 2 & 3 \\ & 0 & -1 & 1 & 0 & 1 & 2 & 1 \\ \hline \mathbf{x} = & 3 & 0 & 0 & 4 & 1 & 0 & \end{array}$$

Do tabulky jsme úplně dolů navíc napsali odpovídající bázové řešení $\mathbf{x}$. Je vidět, že $\mathbf{c}^T\mathbf{x} = 0$ a tedy aktuální hodnota kritéria 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 kritériu neprojeví a kritérium se změní o c_2x_2. Kritérium tedy stoupne nebo zůstane stejné, protože $c_2 = 1 > 0$. $\square$

Pokud v některém sloupci j platí $c_j \leq 0$ a $a_{ij} \leq 0$ pro všechna i, pak můžeme proměnnou x_j libovolně zvětšovat (viz 12.1.3) a úloha je tedy neomezená (její optimum se blíží $-\infty$).

12.2 Základní algoritmus

Spojením popsaných stavebních kamenů dostaneme iteraci simplexové metody. 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 s těmito vlastnostmi:

- podmnožina sloupců $\mathbf{A}$ tvoří standardní bázi,
- bázové řešení odpovídající této bázi je přípustné, tj. $\mathbf{b} \geq \mathbf{0}$,
- složky vektoru $\mathbf{c}$ v bázových sloupcích jsou nulové.

Iteraci se provede ve třech krocích:

1. Vyber index j podle znaménka c_j a index i podle podmínek (12.2).
2. Udělej ekvivalentní úpravu tabulky $[\,\mathbf{A} \,|\, \mathbf{b}\,]$ okolo pivotu (i, j).
3. Udělej ekvivalentní úpravu účelového řádku, která vynuluje c_j v novém bázovém sloupci j.

Požadavky na výběr indexů (i, j) nemusí tyto indexy určit jednoznačně, neboť může být více sloupců j s vhodným znaménkem c_j a více řádků i může splňovat podmínky (12.2). Algoritmus, kterým se vybírá jediný pivot z několika možností, se nazývá **pivotové pravidlo**. Možné (ale nikoliv jediné možné) pivotové pravidlo je následující:

$$j \in \operatorname*{argmin}_{j} c_j, \qquad i \in \operatorname*{argmin}_{i \,|\, a_{ij} > 0} \frac{b_i}{a_{ij}} \tag{12.5}$$

(ve druhém vzorci minimalizujeme přes všechna i splňující $a_{ij} > 0$). Rozmyslete dobře, že druhý vzorec vskutku zajistí splnění podmínek (12.2)! Předpis (12.5) stále nemusí dát jediné (i, j), pokud množiny argmin mají více než jeden prvek – v tom případě vezmeme libovolný z nich.

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 (kritérium nelze zlepšit a jsme v optimu).
- Ve všech sloupcích s $c_j < 0$ je $a_{ij} \leq 0$ pro všechna i (úloha je neomezená). Poznamenejme, že pro neomezenost stačí, když již v *jediném* sloupci je $c_j < 0$ a $a_{ij} \leq 0$ pro všechna i.

Zřídka se algoritmus může dostat do stavu, kdy do nekonečna prochází stále stejnou množinu 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 anticyklící pravidlo*), což ale podrobněji popisovat nebudeme.

Příklad 12.6. Vyřešte simplexovou metodou:

$$\begin{array}{rrcl}
\min & -6x_1 - 8x_2 - 5x_3 - 9x_4 & & \\
\text{za podmínek} & 2x_1 + x_2 + x_3 + 3x_4 + x_5 & = & 5 \\
& x_1 + 3x_2 + x_3 + 2x_4 + x_6 & = & 3 \\
& x_1, x_2, x_3, x_4, x_5, x_6 & \geq & 0
\end{array}$$

Výchozí tabulka je

$$\begin{array}{rrrrrr|r}
-6 & -8 & -5 & -9 & 0 & 0 & 0 \\
\hline
2 & 1 & 1 & 3 & 1 & 0 & 5 \\
1 & 3 & 1 & \boxed{2} & 0 & 1 & 3
\end{array}$$

Účelový řádek budeme nazývat nultý, ostatní pak prvý, druhý atd. Počáteční báze je $L = \{5, 6\}$.

Vybereme sloupec, který nově vstoupí do báze. To může být libovolný sloupec, který má v nultém řádku záporné číslo. Je rozumné vzít největší takové číslo, zde -9, tedy sloupec 4.

Protože do báze chceme přidat sloupec 4, musí některý ze sloupců 5 a 6 z báze ven. Jeho index získáme porovnáním čísel $\frac{5}{3}$ a $\frac{3}{2}$ (čitatel je vždy vpravo, jmenovatel je vždy ve sloupci, který má přijít do báze; uvažujeme ale jen podíly s kladným jmenovatelem): vybereme to nejmenší z nich. Protože $\frac{5}{3} > \frac{3}{2}$, pivot bude v řádku 2. Všimněte si, že přes stávající standardní bázi řádek 2 odpovídá sloupci 2 – tento sloupec tedy půjde z báze ven.

Výsledný pivot (který je v takto nalezeném řádku a sloupci) je označen tučně. Na základě něj spočítáme novou tabulku. To uděláme ekvivalentními řádkovými úpravami, kterými musíme dosáhnout toho, že:

- z pivotu se stane jednička,
- nad i pod pivotem budou nuly, a to včetně nultého řádku.

Jediný způsob, jak toho dosáhnout, je pomocí těchto dvou úprav:

- přičítat vhodné násobky pivotového řádku k ostatním řádkům (tedy k ničemu nikdy nepřičítáme násobky jiného řádku než pivotového)
- samotný pivotový řádek dělit kladným číslem

Tedy k nultému řádku přičteme $\frac{9}{2}$ druhého řádku, k prvnímu řádku přičteme $-\frac{3}{2}$ druhého řádku, a druhý řádek vydělíme dvěma:

$$\begin{array}{rrrrrr|r}
-1.5 & 5.5 & -0.5 & 0 & 0 & 4.5 & 13.5 \\
\hline
\boxed{0.5} & -3.5 & -0.5 & 0 & 1 & -1.5 & 0.5 \\
0.5 & 1.5 & 0.5 & 1 & 0 & 0.5 & 1.5
\end{array}$$

Všimněme si, že vše je v pořádku: v nové tabulce máme opět standardní bázi (sloupce 5 a 4), nad ní máme v nultém řádku nuly, a čísla v nejvíce pravém sloupci jsou nezáporná. V políčku vpravo nahoře máme aktuální hodnotu kritéria, 13.5.

Nový pivot je vyznačen tučně. Další krok je zde:

$$\begin{array}{cccccc|c} 0 & -5 & -2 & 0 & 3 & 0 & 15 \\ \hline 1 & -7 & -1 & 0 & 2 & -3 & 1 \\ 0 & \boxed{5} & 1 & 1 & -1 & 2 & 1 \end{array}$$

Další krok:

$$\begin{array}{cccccc|c} 0 & 0 & -1 & 1 & 2 & 2 & 16 \\ \hline 1 & 0 & 0.4 & 1.4 & 0.6 & -0.2 & 2.4 \\ 0 & 1 & \boxed{0.2} & 0.2 & -0.2 & 0.4 & 0.2 \end{array}$$

Další krok:

$$\begin{array}{cccccc|c} 0 & 5 & 0 & 2 & 1 & 4 & 17 \\ \hline 1 & -2 & 0 & 1 & 1 & -1 & 2 \\ 0 & 5 & 1 & 1 & -1 & 2 & 1 \end{array}$$

Protože všechna čísla v účelovém řádku jsou nezáporná, končíme. Původní LP má optimální řešení, které má hodnotu -17 a nastává v bodě $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6) = (2, 0, 1, 0, 0, 0)$.

Pro kontrolu si dosaďte řešení do původního zadání a zkontrolujte, že řešení je přípustné a že hodnota kritéria je -17. □

Příklad 12.7. Vyřešte simplexovou metodou:

$$\begin{aligned} \min \quad & -2x_1 + 6x_2 + x_3 \\ \text{za podmínek} \quad & -x_1 - x_2 - x_3 + x_4 + = 2 \\ & 2x_1 - x_2 - 2x_3 + + x_5 = 1 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0 \end{aligned}$$

Výchozí tabulka je

$$\begin{array}{ccccc|c} -2 & 6 & 1 & 0 & 0 & 0 \\ \hline -1 & -1 & -1 & 1 & 0 & 2 \\ \boxed{2} & -1 & -2 & 0 & 1 & 1 \end{array}$$

Pivot je vyznačen tučně. Druhá tabulka je zde:

$$\begin{array}{ccccc|c} 0 & 5 & -1 & 0 & 1 & 1 \\ \hline 0 & -1.5 & -2 & 1 & 0.5 & 2.5 \\ 1 & -0.5 & -1 & 0 & 0.5 & 0.5 \end{array}$$

Podle nultého řádku by další pivot měl být ve třetím sloupci. Ale v něm jsou čísla v prvém a druhém řádku záporná. Tedy úloha je neomezená (to bylo ostatně patrné hned ze zadání). V nové tabulce je jasně vidět, že můžeme zvětšovat x_3 libovolně, což bude kompenzováno příslušným nárůstem x_1 a x_4. Jelikož x_1 a x_4 nejsou v kritériu, jejich změny se na něm neprojeví a jediný vliv na kritérium bude mít x_3, které ho bude libovolně zmenšovat. □

12.3 Inicializace algoritmu

Na začátku algoritmu musí být množina přípustných řešení zadána ve tvaru

$$X = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}, \tag{12.6}$$

kde matice $\mathbf{A}$ obsahuje standardní bázi a $\mathbf{b} \geq \mathbf{0}$. Pokud toto není splněno, nemůžeme základní algoritmus přímo aplikovat. Ukážeme, jak lze každou úlohu převést na tento tvar.

Pokud má množina přípustných řešení tvar $\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$ a platí $\mathbf{b} \geq \mathbf{0}$, převod je snadný: přidáme slackové proměnné $\mathbf{u} \geq \mathbf{0}$ a omezení převedeme na tvar $\mathbf{A}\mathbf{x}+\mathbf{u} = \mathbf{b}$. Protože $\mathbf{A}\mathbf{x} + \mathbf{u} = [\,\mathbf{A} \,|\, \mathbf{I}\,] \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix}$, tvoří sloupce příslušné proměnným $\mathbf{u}$ standardní bázi.

Příklad 12.8. Vyřešte simplexovou metodou:

$$\begin{aligned}
\max \quad & 3x_1 + x_2 + 3x_3 \\
\text{za podmínek} \quad & 2x_1 + x_2 + x_3 \leq 2 \\
& x_1 + 2x_2 + 3x_3 \leq 5 \\
& 2x_1 + 2x_2 + x_3 \leq 6 \\
& x_1, x_2, x_3 \geq 0
\end{aligned}$$

Přidáme slackové proměnné $u_1, u_2, u_3 \geq 0$, abychom omezení uvedli do tvaru rovností:

$$\begin{aligned}
\max \quad & 3x_1 + x_2 + 3x_3 \\
\text{za podmínek} \quad & 2x_1 + x_2 + x_3 + u_1 = 2 \\
& x_1 + 2x_2 + 3x_3 + u_2 = 5 \\
& 2x_1 + 2x_2 + x_3 + u_3 = 6 \\
& x_1, x_2, x_3, u_1, u_2, u_3 \geq 0
\end{aligned}$$

Protože hledáme maximum (a ne minimum, jak jsme zvyklí), můžeme buď napsat do tabulky hodnoty $\mathbf{c}$ s opačným znaménkem, nebo vybírat největší (a ne nejmenší) hodnotu v účelovém řádku. Zvolíme duhou možnost. Zde jsou kroky metody:

$$\begin{array}{rrrrrr|r}
3 & 1 & 3 & 0 & 0 & 0 & 0 \\
\hline
\boxed{2} & 1 & 1 & 1 & 0 & 0 & 2 \\
1 & 2 & 3 & 0 & 1 & 0 & 5 \\
2 & 2 & 1 & 0 & 0 & 1 & 6
\end{array}$$

$$\begin{array}{rrrrrr|r}
0 & -0.5 & 1.5 & -1.5 & 0 & 0 & -3 \\
\hline
1 & 0.5 & 0.5 & 0.5 & 0 & 0 & 1 \\
0 & 1.5 & \boxed{2.5} & -0.5 & 1 & 0 & 4 \\
0 & 1 & 0 & -1 & 0 & 1 & 4
\end{array}$$

$$\begin{array}{rrrrrr|r}
0 & -1.4 & 0 & -1.2 & -0.6 & 0 & -5.4 \\
\hline
1 & 0.2 & 0 & 0.6 & -0.2 & 0 & 0.2 \\
0 & 0.6 & 1 & -0.2 & 0.4 & 0 & 1.6 \\
0 & 1 & 0 & -1 & 0 & 1 & 4
\end{array}$$

Úloha má optimální řešení s hodnotou 5.4 v bodě $\mathbf{x} = (x_1, x_2, x_3) = (0.2, 0, 1.6)$ (ověřte v původním zadání!). Hodnota slackových proměnných je $(u_1, u_2, u_3) = (0, 0, 4)$. □

Pokud jsou naše omezení zadána v obecném tvaru, operacemi z §11.1 je lze vždy převést do tvaru (12.6), kde ale nemusí být patrná standardní báze a nemusí platit $\mathbf{b} \geq \mathbf{0}$. Vynásobením některých řádků záporným číslem snadno zajistíme $\mathbf{b} \geq \mathbf{0}$. Matice $\mathbf{A}$ ale nemusí obsahovat standardní bázi. Máme dokonce vážnější problém: není vůbec jasné, zda množina (12.6) je neprázdná, tj. zda je úloha přípustná. V tomto případě nejdříve vyřešíme *pomocnou úlohu* LP, která najde *nějaké* (ne nutně optimální) přípustné řešení. Z něj pak lze získat kýženou standardní bázi. Pomocná úloha zní

$$\min\{\, \mathbf{1}^T\mathbf{u} \mid \mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0},\ \mathbf{u} \geq \mathbf{0}\,\} \tag{12.7}$$

a má simplexovou tabulku

$$\left[\begin{array}{cc|c} \mathbf{0}^T & \mathbf{1}^T & 0 \\ \hline \mathbf{A} & \mathbf{I} & \mathbf{b} \end{array}\right].$$

Je zřejmé, že pomocná úloha má optimum rovné 0 právě tehdy, když $\mathbf{u} = \mathbf{0}$. 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 množina (12.6) je prázdná a původní úloha je nepřípustná.
- Pokud je optimum rovno 0, pak $\mathbf{u} = \mathbf{0}$ a množina (12.6) je neprázdná. Po skončení simplexového algoritmu typicky budou proměnné $\mathbf{u}$ nebázové a tedy mezi sloupci příslušnými proměnným $\mathbf{x}$ bude existovat standardní báze. Kvůli degeneraci se ale zřídka může stát, že některé z proměnných $\mathbf{u}$ budou na konci algoritmu bázové – pak je nutno udělat dodatečné úpravy kolem pivotů ve sloupcích příslušných bázovým proměnným $\mathbf{u}$, abychom tyto proměnné dostali z báze ven.

Nalezení nějakého přípustného řešení v pomocné úloze se nazývá **první fáze** a řešení původní úlohy pak **druhá fáze** algoritmu, mluvíme tedy o **dvoufázové simplexové metodě**.

Příklad 12.9. Řešte

$$\begin{array}{rrcrcrcl} \min & -20x_1 & - & 30x_2 & - & 40x_3 & & \\ \text{za podmínek} & 3x_1 & + & 2x_2 & + & x_3 & = & 10 \\ & x_1 & + & 2x_2 & + & 2x_3 & = & 15 \\ & & & & & x_1, x_2, x_3 & \geq & 0 \end{array}$$

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

$$\begin{array}{rl} \min & u_1 + u_2 \\ \text{za podmínek} & 3x_1 + 2x_2 + x_3 + u_1 \phantom{{}+u_2} = 10 \\ & x_1 + 2x_2 + 2x_3 \phantom{{}+u_1} + u_2 = 15 \\ & x_1, x_2, x_3, u_1, u_2 \geq 0 \end{array}$$

s tabulkou

$$\begin{array}{ccccc|c} 0 & 0 & 0 & 1 & 1 & 0 \\ \hline 3 & 2 & 1 & 1 & 0 & 10 \\ 1 & 2 & 2 & 0 & 1 & 15 \end{array}$$

Sloupce nad přidanými proměnnými tvoří standardní bázi, můžeme tedy na pomocnou úlohu pustit základní simplexový algoritmus. Po vynulování ceny nad bázovými proměnnými budou kroky algoritmu vypadat takto:

$$\begin{array}{rrrrr|r}
-4 & -4 & -3 & 0 & 0 & -25 \\ \hline
3 & \boxed{2} & 1 & 1 & 0 & 10 \\
1 & 2 & 2 & 0 & 1 & 15
\end{array}$$

$$\begin{array}{rrrrr|r}
2 & 0 & -1 & 2 & 0 & -5 \\ \hline
1.5 & 1 & 0.5 & 0.5 & 0 & 5 \\
-2 & 0 & \boxed{1} & -1 & 1 & 5
\end{array}$$

$$\begin{array}{rrrrr|r}
0 & 0 & 0 & 1 & 1 & 0 \\ \hline
2.5 & 1 & 0 & 1 & -0.5 & 2.5 \\
-2 & 0 & 1 & -1 & 1 & 5
\end{array}$$

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 fázi (řešení původní úlohy) s počáteční tabulkou

$$\begin{array}{rrr|r}
-20 & -30 & -40 & 0 \\ \hline
2.5 & 1 & 0 & 2.5 \\
-2 & 0 & 1 & 5
\end{array}$$

□

12.4 Cvičení

12.1. Zapište lineární program

$$\begin{array}{rl}
\min & -x_1 \qquad\qquad\quad - \ x_4 - 3x_5 \\
\text{za podmínek} & 2x_1 \qquad\qquad\quad + \ x_4 + \ x_5 + x_6 = 2 \\
& -x_1 + x_2 \qquad + 2x_4 + 3x_5 \qquad = 4 \\
& 2x_1 \qquad\quad + x_3 + 2x_4 - \ x_5 \qquad = 6 \\
& \qquad\qquad x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
\end{array}$$

do simplexové tabulky. Předpokládejte, že aktuální báze je tvořena sloupci 2, 3, 6 a hodnota kritéria v aktuálním bázovém řešení je nula.

a) Jaké je aktuální bázové řešení?
b) Je toto bázové řešení přípustné či degenerované?
c) Pokud je to možné, udělejte jeden krok simplexového algoritmu, napište simplexovou tabulku po tomto kroku. Pokud to možné není, vysvětlete proč.

12.2. Vyřešte simplexovou metodou:

$$\begin{array}{rrcrcrl}
\max & 2x_1 & - & x_2 & - & 3x_3 & \\
\text{za podmínek} & -2x_1 & - & x_2 & + & x_3 & \leq 2 \\
& -x_1 & + & 2x_2 & - & 3x_3 & \leq 5 \\
& -2x_1 & - & 4x_2 & + & x_3 & \leq 6 \\
& & & & x_1, x_2, & x_3 & \geq 0
\end{array}$$

12.3. Vyřešte simplexovou metodou (navzdory tomu, že lze řešit úvahou):

$$\begin{aligned} \max \quad & 6x_1 + 9x_2 + 5x_3 + 9x_4 \\ \text{za podmínek} \quad & x_1 + x_2 + x_3 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

12.4. Úloha (12.3) má více než jedno optimální řešení. Jak se to projeví v simplexové tabulce? Můžeme udělat výčet všech optimálních bázových řešení?

Kapitola 13

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.

13.1 Konstrukce duální úlohy

K úloze LP v obecném tvaru se duální úloha získá dle tohoto postupu:

$$
\begin{array}{rlrll}
\min & \sum_{j\in J} c_j x_j & \max & \sum_{i\in I} y_i b_i & \\
\text{za podm.} & \sum_{j\in J} a_{ij}x_j = b_i & \text{za podm.} & y_i \in \mathbb{R}, & i\in I_0 \\
& \sum_{j\in J} a_{ij}x_j \geq b_i & & y_i \geq 0\,, & i\in I_+ \\
& \sum_{j\in J} a_{ij}x_j \leq b_i & & y_i \leq 0\,, & i\in I_- \\
& x_j \in \mathbb{R} & & \sum_{i\in I} y_i a_{ij} = c_j, & j\in J_0 \\
& x_j \geq 0 & & \sum_{i\in I} y_i a_{ij} \leq c_j, & j\in J_+ \\
& x_j \leq 0 & & \sum_{i\in I} y_i a_{ij} \geq c_j, & j\in J_-
\end{array}
$$

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, že i-tému primárnímu omezení $\sum_j a_{ij}x_j \geq b_i$ odpovídá duální proměnná $y_i \geq 0$. Tato proměnná je vlastně *Lagrangeův multiplikátor* tohoto omezení. Podobně, j-tá primární proměnná x_j je Lagrangeův multiplikátor j-tého duálního omezení $\sum_i a_{ij}x_j \leq c_j$.

Pro speciální tvary LP se dvojice duálních úloh přehledněji napíše v maticové formě. Např. pro $I_0 = I_- = J_0 = J_- = \emptyset$ obdržíme

$$
\begin{array}{rlrl}
\min & \mathbf{c}^T\mathbf{x} & \max & \mathbf{y}^T\mathbf{b} \\
\text{za podm.} & \mathbf{A}\mathbf{x} \geq \mathbf{b} & \text{za podm.} & \mathbf{y} \geq \mathbf{0} \\
& \mathbf{x} \geq \mathbf{0} & & \mathbf{y}^T\mathbf{A} \leq \mathbf{c}^T
\end{array}
\tag{13.1}
$$

13.2 Věty o dualitě

Následující věty platí pro obecný tvar LP, ale důkazy uděláme pouze pro speciální tvar (13.1).

Věta 13.1 (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 $\mathbf{y}^T\mathbf{A}\mathbf{x} \geq \mathbf{y}^T\mathbf{b}$. Napíšeme-li tyto dvě nerovnosti za sebe, máme

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{y}^T\mathbf{A}\mathbf{x} \geq \mathbf{y}^T\mathbf{b}. \tag{13.2}$$

□

Věta 13.2 (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 komplementarity:*

$$\textit{Pro každé } i \in I \textit{ platí } y_i = 0 \textit{ nebo } \sum_{j\in J} a_{ij}x_j = b_i. \tag{13.3a}$$

$$\textit{Pro každé } j \in J \textit{ platí } x_j = 0 \textit{ nebo } \sum_{i\in I} a_{ij}y_i = c_j. \tag{13.3b}$$

'Nebo' je zde užito v nevylučovacím smyslu, tj. mohou nastat obě možnosti současně.

Důkaz. Klíčové je si uvědomit (rozmyslete!), že pro libovolné vektory $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ platí

$$\forall i\,(u_i = 0 \text{ nebo } v_i = 0) \quad\Longleftrightarrow\quad \forall i\,(u_i v_i = 0) \quad\Longleftrightarrow\quad \mathbf{u}^T\mathbf{v} = 0.$$

Tedy podmínky (13.3) je možno psát jako

$$\mathbf{y}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = 0 \tag{13.4a}$$

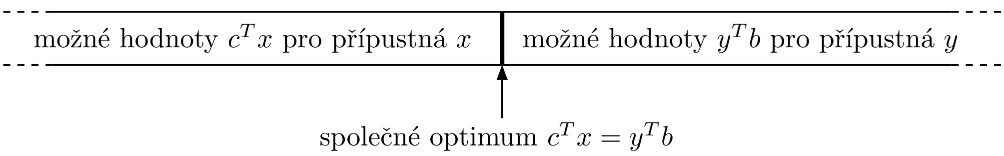
$$(\mathbf{c}^T - \mathbf{y}^T\mathbf{A})\mathbf{x} = 0. \tag{13.4b}$$

Tvrzení $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}$ je ekvivalentní tomu, že obě nerovnosti ve vztahu (13.2) jsou rovnostmi. Pak $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{A}\mathbf{x}$ je ekvivalentní (13.4b) a $\mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{b}$ je ekvivalentní (13.4a). □

Podmínky komplementarity v LP jsou vlastně podmínky komplementarity v KKT-podmínkách pro obecné nelineární optimalizační úlohy.

Věta 13.3 (o silné dualitě). *Primární úloha má optimální řešení právě tehdy, když má duální úloha optimální řešení. Mají-li obě úlohy optimální řešení, platí* $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}$, *kde* $\mathbf{x}$ *a* $\mathbf{y}$ *značí tato optimální řešení.*

Důkaz věty o silné dualitě je obtížný (a vynecháme jej). To není překvapivé, neboť tato věta je jedním z nejhlubších výsledků v lineárním programování. Věty o slabé a silné dualitě mají jasnou interpretaci: pro přípustná $\mathbf{x}$ a $\mathbf{y}$ není hodnota duálního kritéria nikdy větší než hodnota primárního kritéria, a tyto hodnoty se potkají ve společném optimu. Viz obrázek:



Z věty o silné dualitě ihned plyne věta o slabé dualitě (proč?), proto se někdo může ptát, proč větu o slabé dualitě uvádíme. Je to proto, že její důkaz je o mnoho jednodušší.

Dobře si uvědomte, že věta o komplementaritě je slabší než věta o silné dualitě, protože věta o komplementaritě neříká, že rovnost $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}$ vůbec někdy nastane. Uveďme ještě jeden jednoduchý důsledek slabé duality, který je opět slabší než silná dualita.

Důsledek 13.4. *Nechť* $\mathbf{x}$ *je přípustné primární řešení a* $\mathbf{y}$ *je přípustné duální řešení. Nechť* $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}$. *Potom* $\mathbf{x}$ *a* $\mathbf{y}$ *jsou zároveň optimální řešení.*

Důkaz. Pro libovolné primární přípustné řešení $\mathbf{x}'$ plyne z věty o slabé dualitě $\mathbf{y}^T\mathbf{b} \leq \mathbf{c}^T\mathbf{x}'$. Z předpokladu máme $\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{b}$. Z toho plyne $\mathbf{c}^T\mathbf{x} \leq \mathbf{c}^T\mathbf{x}'$. Jelikož toto platí pro *každé* přípustné $\mathbf{x}'$, řešení $\mathbf{x}$ musí být optimální.

Optimalita $\mathbf{y}$ se dokáže symetricky. □

Věta 13.5 (o stínových cenách). *Nechť duální úloha má jediné optimální řešení* $\mathbf{y}^*$. *Označme*

$$f(\mathbf{b}) = \min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\} = \max\{\, \mathbf{y}^T\mathbf{b} \mid \mathbf{y}^T\mathbf{A} \leq \mathbf{c}^T,\ \mathbf{y} \geq \mathbf{0} \,\}$$

optimální hodnotu dvojice duálních úloh jako funkci vektoru $\mathbf{b}$. *Pak je funkce* f *v bodě* $\mathbf{b}$ *diferencovatelná a platí* $f'(\mathbf{b}) = (\mathbf{y}^*)^T$.

Důkaz. Jelikož je optimální řešení $\mathbf{y}^*$ jediné, nabývá se ve vrcholu polyedru přípustných řešení $\{\, \mathbf{y} \in \mathbb{R}^m \mid \mathbf{y}^T\mathbf{A} \leq \mathbf{c}^T,\ \mathbf{y} \geq \mathbf{0} \,\}$. Změníme-li nepatrně $\mathbf{b}$, množina duálních optimálních řešení se nezmění, neboli budeme mít stále jediné optimální řešení ve stejném vrcholu $\mathbf{y}^*$ (tento argument není zcela rigorózní, ale geometricky je dosti názorný). Tedy v malém okolí bodu $\mathbf{b}$ je hodnota optima jednoduše rovna $(\mathbf{y}^*)^T\mathbf{b}$. Derivací toho získáme $f'(\mathbf{b}) = (\mathbf{y}^*)^T$. □

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 $\mathbf{A}\mathbf{x} \geq \mathbf{b}$. Interpretujeme-li naše LP jako optimální výrobní plán (11.3) (pozor, liší se obrácenou nerovností v omezení), pak hodnota y_i^* říká, jak by se náš výdělek zvětšil, kdybychom trochu uvolnili omezení na výrobní zdroje $\mathbf{a}_i^T\mathbf{x} \leq b_i$. Proto se duálním proměnným někdy 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 13.1. Mějme dvojici navzájem duálních úloh LP:

$$\begin{array}{rrcrcrcrl}
 & \min & 2x_1 & + & 5x_2 & + & 6x_3 & = & \mathbf{5.4} \\
\mathbf{3} = & & 2x_1 & + & x_2 & + & 2x_3 & \geq & 3 \\
\mathbf{2.4} = & & x_1 & + & 2x_2 & + & 2x_3 & \geq & 1 \\
\mathbf{3} = & & x_1 & + & 3x_2 & + & x_3 & \geq & 3 \\
\mathbf{-0.6} = & & -x_1 & + & x_2 & - & 2x_3 & \geq & -1 \\
\mathbf{1.2} = & & x_1 & & & & & \geq & 0 \\
\mathbf{0.6} = & & & & x_2 & & & \geq & 0 \\
\mathbf{0} = & & & & & & x_3 & \geq & 0
\end{array}
\qquad
\begin{array}{rcrcrcrcrl}
\max & 3y_1 & + & y_2 & + & 3y_3 & - & y_4 & = & \mathbf{5.4} \\
\mathbf{0.2} = & y_1 & & & & & & & \geq & 0 \\
\mathbf{0} = & & & y_2 & & & & & \geq & 0 \\
\mathbf{1.6} = & & & & & y_3 & & & \geq & 0 \\
\mathbf{0} = & & & & & & & y_4 & \geq & 0 \\
\mathbf{2} = & 2y_1 & + & y_2 & + & y_3 & - & y_4 & \leq & 2 \\
\mathbf{5} = & y_1 & + & 2y_2 & + & 3y_3 & + & y_4 & \leq & 5 \\
\mathbf{2} = & 2y_1 & + & 2y_2 & + & y_3 & - & 2y_4 & \leq & 6
\end{array}$$

Spočetli jsme optimální řešení obou úloh a dosadili tato řešení do kritérií 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 kritérií v optimech jsou napsané tučně před/za rovnítky. Vidíme, že:

- Obě optima se sobě rovnají. Tak to musí být podle věty o silné dualitě.
- Vezmeme-li libovolný řádek (kromě kriteriálního), je na něm alespoň jedno z obou omezení aktivní (tj. platí s rovností). Např. ve druhém řádku je primární omezení $2x_1+x_2+2x_3 \geq 3$ aktivní (protože jeho levá strana je rovna 3) a duální omezení $y_1 \geq 0$ je neaktivní (protože platí $y_1 > 0$). Podle věty o komplementaritě se nemůže stát, že by na některém řádku byly obě rovnosti zároveň neaktivní (mohou být obě ale zároveň aktivní, což zde nenastává, ale stane se to v případě degenerace).
- Např. $y_1 = 0.2$ je stínová cena prvního primárního omezení $2x_1 + x_2 + 2x_3 \geq 3$. 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 změní 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:

$$
\begin{array}{rrrrcr}
\min & 2x_1 + & 5x_2 + & 6x_3 & = & \mathbf{5.402} \\
& 2x_1 + & x_2 + & 2x_3 & \geq & 3.01 \\
& x_1 + & 2x_2 + & 2x_3 & \geq & 1 \\
& x_1 + & 3x_2 + & x_3 & \geq & 3 \\
& -x_1 + & x_2 - & 2x_3 & \geq & -1 \\
& x_1 & & & \geq & 0 \\
& & x_2 & & \geq & 0 \\
& & & x_3 & \geq & 0
\end{array}
\qquad
\begin{array}{rrrrrcr}
\max & 3.01y_1 + & y_2 + & 3y_3 - & y_4 & = & \mathbf{5.402} \\
\mathbf{0.2} = & y_1 & & & & \geq & 0 \\
\mathbf{0} = & & y_2 & & & \geq & 0 \\
\mathbf{1.6} = & & & y_3 & & \geq & 0 \\
\mathbf{0} = & & & & y_4 & \geq & 0 \\
\mathbf{2} = & 2y_1 + & y_2 + & y_3 - & y_4 & \leq & 2 \\
\mathbf{5} = & y_1 + & 2y_2 + & 3y_3 + & y_4 & \leq & 5 \\
\mathbf{2} = & 2y_1 + & 2y_2 + & y_3 - & 2y_4 & \leq & 6
\end{array}
$$

  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

$$5.402 - 5.4 = \frac{\partial f(\mathbf{b})}{\partial b_1} h = y_1 h = 0.2 \cdot 0.01.$$

□

Příklad 13.2. Je dána primární úloha z Příkladu 13.1. Je dáno $\mathbf{x} = (x_1, x_2, x_3) = (1.2, 0.6, 0)$, ale už není dáno duální řešení $\mathbf{y}$. Dokažte bez použití algoritmu na řešení LP, že dané $\mathbf{x}$ je optimální řešení primární úlohy.

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

$$
\begin{aligned}
2y_1 + y_3 &= 2 \\
y_1 + 3y_3 &= 5
\end{aligned}
\tag{13.5}
$$

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 primárního kritéria v bodě $\mathbf{x}$ rovná hodnotě duálního kritéria v bodě $\mathbf{y}$, musejí být $\mathbf{x}$ a $\mathbf{y}$ optimální řešení.

Zdůrazněme, že tento postup nemusí vést vždy k cíli. Když bude mít duální úloha více než jedno řešení, bude mít soustava (13.5) nekonečně mnoho (afinní podprostor) řešení. Mezi nimi sice budou přípustná duální řešení, ale k jejich nalezení budeme potřebovat řešit soustavu rovnic *a nerovnic* (což už není snadné). □

13.2.1 Dovolené kombinace řešitelnosti dvojice duálních úloh

Zopakujme, že pro každou úlohu LP mohou nastat 3 možnosti:

- úloha má optimální řešení
- úloha je neomezená
- úloha je nepřípustná (tj. množina přípustných řešení je prázdná).

Věta 13.6. *Z devíti možností pro dvojici duálních úloh se realizují tyto:*

primární/duální	*má optimum*	*neomezená*	*nepřípustná*
má optimum	*ano*	*ne*	*ne*
neomezená	*ne*	*ne*	*ano*
nepřípustná	*ne*	*ano*	*ano*

Důkaz. Snadno najdeme příklady dvojic duálních úloh, které realizují povolené kombinace. Zbývá dokázat, že zakázané kombinace nemohou nastat.

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).

Pokud je primární resp. úloha neomezená, její optimum je $-\infty$ resp. $+\infty$. Věta o slabé dualitě zakazuje, aby úlohy byly zároveň neomezené, protože pak bychom měli $-\infty \geq +\infty$. □

Často se užívá konvence, že optimum nepřípustné minimalizační úlohy je rovno $+\infty$ a optimum nepřípustné maximalizační úlohy je rovno $-\infty$. To vyplývá z toho, že $\inf \emptyset = +\infty$ a $\sup \emptyset = -\infty$, viz §6.1. Za této konvence je třetí řádek třetí a sloupec ve shodě s větou o slabé dualitě, i když přísně řečeno toto tvrzení z ní nevyplývá (protože věta nic neříká o nepřípustných řešeních).

13.3 Použití duality

- *Návrh nových algoritmů.* Zatímco přímá úloha má m omezení a n proměnných, duální úloha má n omezení a m proměnných.
- *Certifikát optimality.* Předložíme-li přípustná primární a duální řešení taková, že se kritéria rovnají, dokázali jsme optimalitu. Pro úlohy s mnoha omezeními/proměnnými to může být nejsnadnější důkaz optimality. Někdy lze spočítat optimální duální řešení levně z optimálního primárního, viz Příklad 13.2.
- Dualita umožňuje *vhled* do řešeného problému, často velmi netriviální. Dá se říci, že abychom jakékoli úloze porozuměli do hloubky, musíme pochopit jak její primární tak duální formulaci.
- *Citlivostní analýza, stínové ceny.*

Příklad 13.3. Mějme úlohu

$$\min\{\, \mathbf{a}^T\mathbf{x} \mid \mathbf{1}^T\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0} \,\} = \min \left\{ \sum_{i=1}^n a_i x_i \;\middle|\; \sum_{i=1}^n x_i = 1,\ x_i \geq 0 \right\},$$

kde $\mathbf{a} = (a_1, \ldots, a_n)$ je dáno a optimalizuje se přes $\mathbf{x} = (x_1, \ldots, x_n)$. Najděte elementární úvahou hodnotu optima, napište duální úlohu. Vysvětlete, co v dané úloze znamenají věty o silné dualitě a komplementaritě.

Optimální hodnota je $\min_{i=1}^n a_i$, tedy nejmenší z čísel a_i. Dosahuje se ve vektoru $\mathbf{x}$ jehož všechny složky jsou nulové kromě složek příslušných minimálnímu a_i. To je jasné, protože je nejvýhodnější soustředit všechnu 'váhu' rozdělení $\mathbf{x}$ do nejmenšího prvku. Pokud je více minimálních prvků a_i, optimální vektor $\mathbf{x}$ není dán jednoznačně. Např. pro $\mathbf{a} = (1, 3, 1, 2)$ 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

$$\max\{\, y \in \mathbb{R} \mid y\mathbf{1} \leq \mathbf{a} \,\} = \max\{\, y \in \mathbb{R} \mid y \leq a_i,\ i = 1, \ldots, n \,\}.$$

Tato úloha má jasný význam: hledá se největší číslo y, které je menší než všechna čísla a_i. Takové číslo y se rovná minimu z čísel a_i.

Význam silné duality je jasný: hodnoty primárního i duálního optima jsou si rovny.

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 $x_i \geq 0$, $y \leq a_i$ je splňuje právě tehdy, když je splněna aspoň jedna z rovností $x_i = 0$, $y = a_i$. To znamená:

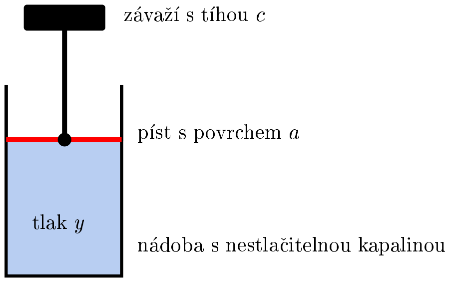
- Pokud je v duálu $y < a_i$, v primáru musí být $x_i = 0$. To je ale jasné, protože $y < a_i$ znamená, že a_i není nejmenší ze složek vektoru $\mathbf{a}$ a tudíž v primáru by byla hloupost mu příŕadit nenulovou váhu x_i.
- Obráceně, pokud je v primáru $x_i > 0$, musí být v duálu $y = a_i$. To je jasné, protože pokud jsme v primáru přiřadili číslu a_i nenulovou váhu, musí být nejmenší. □

13.4 Ilustrace duality na hydraulickém počítači

(Tato část je nepovinná, doporučujeme ale projít.)

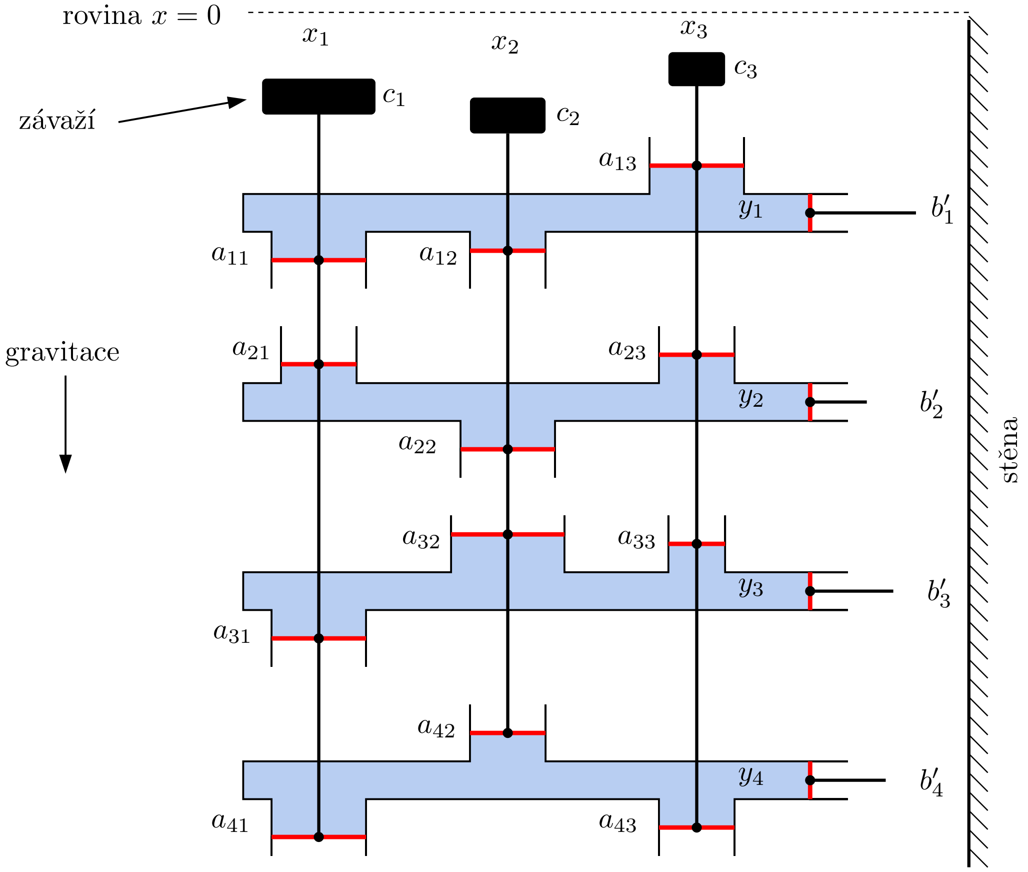
Uvažujme fyzikální systém ('analogový počítač'), který sestává z nádob s nestlačitelnou kapalinou uzavřených písty a ze závaží. K pochopení jeho činnosti budeme potřebovat tyto známé poučky z hydrostatiky:

- Objem kapaliny v uzavřené nádobě je při libovolném tlaku stejný.
- Tlak v kapalině je při rovnováze všude stejný.
- Nechť y je tlak v nádobě s kapalinou uzavřené pístem. Nechť píst má povrch a a působí na něj síla c (viz obrázek). Pak $c = ay$.



Obrázek 13.1 ukazuje celý stroj, v němž

- a_{ij} = povrch svislého pístu v nádobě i spojeného se závažím j (pro $a_{ij} > 0$ je píst nahoře, pro $a_{ij} < 0$ je píst dole).
- x_j = výška závaží j (měřeno směrem dolů od roviny nulové výšky)
- b'_i = šířka mezery mezi tyčí a stěnou. Při $\mathbf{x} = \mathbf{0}$ platí $b'_i = b_i$.
- c_j = tíha závaží j
- Vodorovné písty mají povrch jednotkový.



Obrázek 13.1: Hydraulický počítač řešící primární i duální úlohu LP.

Ze zachování objemu kapaliny v nádobě i plyne $b'_i = b_i - a_{i1}x_1 - \cdots - a_{in}x_n$. Protože vodorovné tyče nemohou projít stěnou, je vždy $b'_i \geq 0$. Tedy $\mathbf{b}' = \mathbf{b} - \mathbf{A}\mathbf{x} \geq \mathbf{0}$, tedy $\mathbf{A}\mathbf{x} \leq \mathbf{b}$.

Potenciální energie závaží j je $-c_j x_j$. Libovolný statický systém v rovnováze zaujme stav s nejnižší potenciální energií. Proto se závaží ustálí v takových výškách, že jejich celková potenciální energie bude minimální, neboli $c_1x_1 + \cdots + c_nx_n = \mathbf{c}^T\mathbf{x}$ bude maximální. Tedy stroj řeší lineární program

$$\max\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{R}^n \,\}.$$

Jeho duál je

$$\min\{\, \mathbf{y}^T\mathbf{b} \mid \mathbf{y}^T\mathbf{A} = \mathbf{c}^T,\ \mathbf{y} \ge \mathbf{0} \,\}.$$

Tento duální LP jsme našli *čistě formálně* podle postupu o konstrukci duální úlohy. To nám ale vůbec neříká, jaký má duál vztah k našemu stroji. Pokusme se tento vztah odhalit. Klíčové pro objevení tohoto vztahu je přiřadit duální proměnné y_i význam *tlaku* v nádobě i (všimněte si, že v primární úloze tlak vůbec nevystupuje). Teď dokážeme interpretovat duální LP a jeho vztah k primárnímu LP daný větami o dualitě (vynecháváme větu o slabé dualitě):

- Jelikož stěna působí silou vždy od sebe, musí být tlak v y_i v nádrži i nezáporný. To dá duální omezení $\mathbf{y} \ge \mathbf{0}$.
- Protože povrch vodorovných pístů je jednotkový, tlak y_i se rovná síle vodorovné tyče i na stěnu. Rovnováha sil pro svislou tyč j bude $a_{1j}y_1 + \cdots + a_{mj}y_m = c_j$, což je duální omezení $\mathbf{y}^T\mathbf{A} = \mathbf{c}^T$.
- Dle *věty o komplementaritě* v ustáleném stavu platí buď $b'_i = b_i - a_{i1}x_1 - \cdots - a_{in}x_n = 0$ nebo $y_i = 0$, pro každé i. Ale to je jasné, protože když se některá vodorovná tyč nedotýká stěny, musí být tlak v příslušné nádobě nulový.
- Dle *věty o silné dualitě* je v ustáleném stavu duální kritérium $\mathbf{y}^T\mathbf{b} = y_1b_1 + \cdots + y_mb_m$ minimální. Proč to tak je? Potenciální energie všech závaží je rovna práci, nutné na jejich vyzdvižení do roviny $x = 0$. Tato práce se dá vykonat buď přímo zdvihnutím závaží (což odpovídá primárnímu kritériu $\mathbf{c}^T\mathbf{x}$) nebo odtlačením vodorovných tyčí od stěny do vzdáleností b_i. Ukážeme, že druhý způsob odpovídá duálnímu kritériu. Zafixujeme-li všechny vodorovné tyče kromě jediné tyče i, při odtlačování tyče i se síla, kterou tlačíme na tyč, nemění (promyslete!). Tedy vykonáme práci y_ib_i. Když takto odtlačíme od stěny postupně všechny tyče, vykonáme práci $\mathbf{y}^T\mathbf{b}$.
- Věta o *stínových cenách* říká, že se změnou b_i se optimum mění tím více, čím je větší tlak y_i. To je ale jasné, protože čím je větší tlak y_i, tím větší práce je třeba na odtlačení tyče od stěny do vzdálenosti b_i.

Zdůrazněme, že tyto úvahy *nedokazují* žádnou ze tří vět o dualitě. Předpokládáme totiž platnost fyzikálních zákonů, které ale nelze matematicky dokázat, lze je pouze experimentálně pozorovat. Skutečnost, že z chování stroje 'vyplývá' např. věta o silné dualitě, není matematický důkaz – ten je totiž čistou logickou dedukcí a žádné fyzikální zákony nepředpokládá.

Tím, že se nám podařilo pochopit význam duální úlohy ve stroji, jsme se o fyzice našeho stroje dozvěděli něco nového – tedy, že se dá podmínka rovnováhy formulovat pomocí tlaků v nádobách. Toho bychom si nejspíše nevšimli, kdybychom se nezabývali duální úlohou. Tak je to u každého reálného systému (fyzikálního, ekonomického, ...) popsaného lineárním programem: abychom systém pochopili do hloubky, je třeba přiřadit *význam* nejen primární úloze, ale i duální úloze a větám o dualitě.

13.5 Cvičení

13.1. Ukažte pro dvojici úloh LP v §13.1, 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.

13.2. Napište duál a podmínky komplementarity pro úlohy (11.6) a (11.8).

13.3. Napište duální úlohy a podmínky komplementarity k následujícím úlohám. Výsledek vždy co nejvíce zjednodušte příp. převeďte z maticové do skalární formy, je-li skalární forma výstižnější. Pro každou úlohu interpretujte silnou duality (tj. úvahou odvoďte, jaká je optimální hodnota duální úlohy, a tato musí být stejná jako optimální hodnota primární úlohy) a podmínky komplementarity, podobně jako v Příkladu 13.3.

a) Všechny úlohy ze Cvičení 11.2.
b) $\min_{x\in\mathbb{R}} \max_{i=1}^{n} |a_i - x|$ (střed intervalu)
c) $(\star)$ $\min_{x\in\mathbb{R}} \sum_{i=1}^{n} |a_i - x|$ (medián)
d) Úloha LP vzniklá ve Cvičení 11.6.
e) $(\star)$ Příklad 11.4.

13.4. Dokažte bez užití algoritmu na řešení LP, že $\mathbf{x} = (1, 1, 1, 1)$ je optimální řešení úlohy

$$\begin{aligned} \min \quad & \begin{bmatrix} 47 & 93 & 17 & -93 \end{bmatrix}^T \mathbf{x} \\ \text{za podm.} \quad & \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \mathbf{x} \le \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix} \end{aligned}$$