Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Lukáš Adam

Konvexní analýza v optimalizaci

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Doc. RNDr. Jiří Spurný, Ph.D., katedra matematické analýzy
Studijní program: Obecná matematika

2008

Chtěl bych tímto poděkovat Doc. RNDr. Jiřímu Spurnému, Ph.D. za spolehlivé a poučné vedení této práce, konstruktivní připomínky a velké množství času, které věnoval mé práci.

Prohlašuji, že jsem svou bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 25.5.2008 Lukáš Adam

Obsah

Název práce: Konvexní analýza v optimalizaci
Autor: Lukáš Adam
Vedoucí bakalářské práce: Doc. RNDr. Jiří Spurný, Ph.D., katedra matematické analýzy
E-mail vedoucího: spurny@karlin.mff.cuni.cz

Abstrakt: V předložené práci studujeme v první kapitole oddělovací věty a dále pak jejich aplikace. V druhé kapitole se potom zabývám variačními principy, nejdříve Ekelandovým, postupně přecházím k Borwein-Preissovu a nakonec k Lindenstrauss-Preiss-Tišerovu variačnímu principu. Na závěr se věnuji Banachově větě o kontrakci a její souvislosti s úplností metrického prostoru.

Klíčová slova: Oddělování, variační principy, úplnost.

Title: Convex analysis in optimalization
Author: Lukas Adam
Supervisor: Doc. RNDr. Jiří Spurný, Ph.D.
Supervisor's e-mail address: spurny@karlin.mff.cuni.cz

Abstract: In the present work we study in the first chapter various separation theorems and their applications. In the second chapter we study variational principles, at first Ekeland's, then Borwein-Preiss' and finally Lindenstrauss-Preiss-Tiser's variational principle. At the end we study the fix point theorem and its connection to completeness of a metric space.

Keywords: Separation theorems, variational principles, completeness.

Úvod

Práce je rozdělena do dvou kapitol. První kapitola se zabývá oddělováním množin a jejich aplikacemi na řešení lineárních rovnic, druhá pak Ekelandovým variačním principem, jeho aplikacemi a zobecněními. V první kapitole uvádím různé varianty vět o oddělování. V celé této kapitole pracuji pouze s reálnými konečně rozměrnými vektorovými prostory.

V druhé kapitole pracuji s obecnými metrickými prostory. Nejdříve uvádím důkaz Ekelandova variačního principu a poté několik jeho přímých důsledků. Jeho užitečnost je doložena důkazy několika alterativ pro řešení soustav lineárních rovnic. Dokazuji i Stiemkeho alternativu, kterou jsem dokázal již v první kapitole.

V další části dokazuji zobecnění Ekelandova variačního principu, nejdříve Borwein-Preissův variační princip a později dokonce i Lindenstrauss-Preiss-Tišerův variační princip.

Nakonec ukazuji, že Ekelandův variační princip je úzce spjatý s úplností metrického prostoru. Platí totiž, že metrický prostor je úplný tehdy a jen tehdy, platí-li Ekelandův variační princip. Na závěr se obracím k Banachově větě o kontrakci, kterou pomocí Ekelandova variačního principu dokazuji. Zároveň dokazuji, že na rozdíl Ekelandova variačního principu není Banachova věta o kontrakci ekvivalentní s úplností metrického prostoru. Uvádím totiž příklad metrického prostoru který není úplný, ale každá kontrakce na něm má pevný bod.

Kapitola 1

Oddělování

V celé této kapitole budeme pracovat s reálnými konečně dimenzionálními prostory $\mathbb{R}^n$. Dále budeme uvažovat euklidovskou metriku generovanou skalárním součinem, který budeme značit $\langle \cdot, \cdot \rangle$. Položme $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Jsou-li $x, y \in \mathbb{R}^n$, pak píšeme $x = (x_1, ..., x_n)$. Zápisem $x > y$ rozumíme, že platí $x_i \geq y_i,\ i = 1, ..., n$ a $x \neq y$. Zápisem $x \gg y$ rozumíme $x_i > y_i,\ i = 1, ..., n$. Označme $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n; x \gg 0\}$ a $\mathbb{R}^n_{+0} = \{x \in \mathbb{R}^n; x > 0\}$. Pokud máme matici A o rozměrech $m \times n$ pak $a_1, ..., a_n$ budou označovat její sloupce.

Je-li $K \subset \mathbb{R}^n$, pak zápisem $\overline{K}$ rozumíme uzávěr K, zápisem K° vnitřek množiny, zápisem ∂K hranici K a konečně zápisem $\mathrm{rint}(K)$ relativní vnitřek K, tedy vnitřek vzhledem k nejmenší afinní nadrovině, která je nadmnožinou K.

Pokud je d metrika nebo pseudometrika, zápisem $B_d(x, r)$ rozumíme otevřenou kouli o středu x a poloměru r, kde měříme pomocí d. Pokud nemůže dojít k omylu, budu používat pouze $B(x, r)$. Zápis S_X pak pro normovaný lineární prostor X označuje jednotkovou sféru prostoru X.

1.1 Oddělovací věty

Věta 1.1.1 (o projekci). *Nechť $K \subset \mathbb{R}^n$ je neprázdná uzavřená konvexní množina a nechť $x \notin K$. Potom existuje právě jedno $y \in K$ splňující*

$$\|x - y\| = \inf\{\|x - z\|; z \in K\}.$$

Tento bod je jednoznačně určen podmínkou $\langle x - y, z - y \rangle \leq 0$ pro všechny $z \in K$.

Důkaz. Definujme $d = \inf\{\|x - z\|; z \in K\}$. Potom existuje posloupnost $\{y_n\}$ bodů v $\mathbb{R}^n$ taková, že $\|x - y_n\| \to d$. Z rovnoběžníkového pravidla dostáváme, že pro každé

$m, n \in \mathbb{N}$ platí

$$\|x - y_n + x - y_m\|^2 + \|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2,$$

z čehož dokážeme, že posloupnost $\{y_n\}$ je cauchyovská. Máme totiž

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\|x - \frac{y_n - y_m}{2}\|^2 \\ &\leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4d^2,\end{aligned}$$

kde poslední člen konverguje k nule pro $m, n \to \infty$.

Z úplnosti $\mathbb{R}^n$ a uzavřenosti K pak plyne $y_n \to y$ pro nějaké $y \in K$, ze spojitosti normy potom dostáváme $\|x - y\| = d$.

Pro důkaz jednoznačnosti předpokládejme, že nějaké $y' \in K$ splňuje $\|x - y'\| = d$. Potom ale platí

$$\|y - y'\|^2 = 2\|x - y\|^2 + 2\|x - y'\|^2 - 4\|x - \frac{y + y'}{2}\|^2 \leq 4d^2 - 4d^2 = 0,$$

a tedy y je určeno jednoznačně a zbývá dokázat druhou část věty.

Ukažme nejdříve, že $\langle x - y, z - y\rangle$ pro všechna $z \in K$. Pro spor předpokládejme, že existuje $z \in K$ splňující $\langle x-y, z-y\rangle > 0$. Definujme kvadratickou funkci $f : \mathbb{R} \to [0, \infty)$ následovně:

$$\begin{aligned}f(\alpha) &= \|x - (\alpha y + (1 - \alpha)z)\|^2 = \|\alpha(x - y) + (1 - \alpha)(x - z)\|^2 \\ &= \langle \alpha(x - y) + (1 - \alpha)(x - z), \alpha(x - y) + (1 - \alpha)(x - z)\rangle \\ &= \alpha^2\|x - y\|^2 + (1 - \alpha)^2\|x - z\|^2 + 2\alpha(1 - \alpha)\langle x - y, x - z\rangle.\end{aligned}$$

Pak

$$f(1) = \|x - y\|^2 = d^2$$

a

$$\begin{aligned}f'(1) &= 2\|x - y\|^2 - 2\langle x - y, x - z\rangle \\ &= 2\|x - y\|^2 - 2\langle x - y, x - y\rangle - 2\langle x - y, y - z\rangle \\ &= 2\langle x - y, z - y\rangle > 0.\end{aligned}$$

Existuje tedy $t < 1$ takové, že $\|x - (ty + (1 - t)z)\| < d$. Z konvexity pak plyne, že $ty + (1 - t)z \in K$, čímž docházíme ke sporu s definicí d.

Dokažme nyní obrácenou implikaci. Nechť $y \in K$ splňuje $\langle x - y, z - y\rangle \leq 0$ pro všechna $z \in K$. Potom pro libovolné $z \in K$ máme

$$\begin{aligned}\|x - z\|^2 &= \|(x - y) - (z - y)\|^2 \\ &= \langle (x - y) - (z - y), (x - y) - (z - y)\rangle \\ &= \|x - y\|^2 + \|z - y\|^2 - 2\langle x - y, z - y\rangle \geq \|x - y\|^2\end{aligned}$$

a tedy $\|x-y\| = \inf\{\|x-z\|; z \in K\}$, čímž je věta dokázána.

$\square$

Věta 1.1.2 (oddělovací věta pro bod a množinu)**.** *Nechť $K \subset \mathbb{R}^n$ je neprázdná konvexní množina a nechť $x \notin \overline{K}$. Potom existuje $\gamma \in \mathbb{R}^n$ splňující*

$$\inf\{\langle\gamma, z\rangle; z \in K\} > \langle\gamma, x\rangle.$$

Důkaz. Protože uzávěr konvexní množiny je konvexní uzavřená množina, platí pro $\overline{K}$ předpoklady věty 1.1.1. Existuje tedy $y \in \overline{K}$ realizující vzdálenost x od $\overline{K}$. Pokud definujeme $\gamma = y - x$, pak pro všechna $z \in \overline{K}$ platí

$$\begin{aligned} \langle x-y, z-y\rangle &\leq 0, \\ \langle -\gamma, z-y\rangle &\leq 0, \\ \langle\gamma, y\rangle &\leq \langle\gamma, z\rangle. \end{aligned}$$

Protože $x \notin \overline{K}$ a $y \in \overline{K}$, platí

$$0 < \|\gamma\|^2 = \langle\gamma, y-x\rangle = \langle\gamma, y\rangle - \langle\gamma, x\rangle,$$

a tedy je

$$\inf\{\langle\gamma, z\rangle; z \in K\} \geq \inf\{\langle\gamma, z\rangle; z \in \overline{K}\} \geq \langle\gamma, y\rangle > \langle\gamma, x\rangle,$$

čímž je důkaz dokončen.

$\square$

Poznámka 1.1.3. Všimněme si, že kvůli ostré nerovnosti nikdy nemůže nastat případ $\gamma = 0$. Toto bude několikrát využito v následujících větách.

Poznámka 1.1.4. Ukažme si geometrickou interpretaci předchozí věty. Vezměme si libovolné pevné $a \in \mathbb{R}^n$ a definujme funkci $f(x) = \langle a, x\rangle,\ x \in \mathbb{R}^n$. Předchozí věta tedy říká, že můžeme najít lineární funkci f na $\mathbb{R}^n$ a $c \in \mathbb{R}$ takové, že afinní nadrovina $\{x \in \mathbb{R}^n; f(x) = c\}$ odděluje bod od množiny. Funkce f je určena pomocí vektoru γ. Hodnotu c si můžeme zvolit jako jakékoli číslo z intervalu

$$(\langle\gamma, x\rangle, \inf\{\langle\gamma, z\rangle, z \in K\}].$$

Věta 1.1.5 (o opěrné nadrovině)**.** *Nechť $K \subset \mathbb{R}^n$ je neprázdná konvexní množina a nechť $x \in \partial K$. Potom existuje $\gamma \in \mathbb{R}^n \backslash \{0\}$ takové, že*

$$\inf\{\langle\gamma, z\rangle; z \in K\} \geq \langle\gamma, x\rangle.$$

Důkaz. Dle předpokladu můžeme najít posloupnost $x_i \to x, x_i \notin \overline{K}$. Postupně užíváme oddělovací větu 1.1.2 na body x_i a množinu K. Pro každé $i \in \mathbb{N}$ tedy existuje $\gamma_i \in \mathbb{R}^n \backslash \{0\}$ splňující

$$\inf\{\langle \gamma_i, z\rangle; z \in K\} > \langle \gamma_i, x_i\rangle.$$

Bez újmy na obecnosti můžeme předpokládat, že tyto nalezené vektory mají jednotkovou normu. Protože $\{\gamma_i; i \in \mathbb{N}\} \subset S_{\mathbb{R}^n}$, můžeme z této posloupnosti vybrat konvergentní podposloupnost $\gamma_{i_k} \to \gamma$.

Pro libovolné $z \in K$ pak dostáváme

$$\langle \gamma, z\rangle = \lim_{k\to\infty} \langle \gamma_{i_k}, z\rangle \geq \lim_{k\to\infty} \langle \gamma_{i_k}, x_{i_k}\rangle = \langle \gamma, x\rangle,$$

čímž je důkaz hotov.

□

Věta 1.1.6 (oddělovací věta pro dvě množiny)**.** *Nechť $A, B \subset \mathbb{R}^n$ jsou neprázdné konvexní množiny splňující $A \cap B = \emptyset$. Potom existuje $\gamma \in \mathbb{R}^n \backslash \{0\}$ takové, že platí*

$$\inf\{\langle \gamma, a\rangle; a \in A\} \geq \sup\{\langle \gamma, b\rangle; b \in B\}.$$

Důkaz. Pokud je alespoň jedna množina jednobodová, věta se redukuje na větu 1.1.5. Předpokládejme tedy, že množiny A, B nejsou jednobodové. Položme $K = A - B$. Protože $A \cap B = \emptyset$, je $K \cap \{0\} = \emptyset$ a ze vztahu $\overline{K} = K \cup \partial K$ plyne, že platí buď $0 \notin \overline{K}$ nebo $0 \in \partial K$.

V prvním případě použijeme větu 1.1.2, v druhém pak 1.1.5. V obou případech vyjde, že můžeme 0 oddělit neostře od množiny K. To znamená, že existuje $\gamma \in \mathbb{R}^n \backslash \{0\}$ takové, že platí

$$\begin{aligned} 0 &\leq \inf\{\langle \gamma, x\rangle; x \in K\} = \inf\{\langle \gamma, a - b\rangle; a \in A, b \in B\} \\ &= \inf\{\langle \gamma, a\rangle - \langle \gamma, b\rangle; a \in A, b \in B\} = \inf\{\langle \gamma, a\rangle; a \in A\} - \sup\{\langle \gamma, b\rangle; b \in B\} \end{aligned}$$

což je přesně požadované tvrzení.

□

1.2 Aplikace oddělovacích vět

Definice 1.2.1. Množinu $K \subset \mathbb{R}^n$ nazveme kužel, pokud je neprázdná a pro všechny $x \in K, t \geq 0$ platí $tx \in K$.

Množinu $K \subset \mathbb{R}^n$ nazveme konvexní kužel, pokud je konvexní a kužel.

Poznámka 1.2.2. Je snadno vidět, že množina $K \subset \mathbb{R}^n$ je konvexní kužel tehdy a jen tehdy, pokud pro libovolné $x, y \in K$ a $t \geq 0$ platí $x + y \in K, tx \in K$.

Věta 1.2.3 (Alternativa pro kužel)**.** *Nechť $K \subset \mathbb{R}^n$ je konvexní kužel. Potom pro každé $x \in \mathbb{R}^n$ platí právě jedno z následujících:*

(i) $x \in \overline{K}$,

(ii) $\exists \gamma \in \mathbb{R}^n \ \forall y \in K : \langle \gamma, y \rangle \geq 0, \langle \gamma, x \rangle < 0$.

Důkaz. Pokud $x \in \overline{K}$, pak $\inf\{\langle \gamma, y \rangle; y \in K\} \leq \langle \gamma, x \rangle$, a tedy nemůže platit druhá možnost.

Předpokládejme, že $x \notin \overline{K}$. Použitím oddělovací věty 1.1.2 dostáváme existenci $\gamma \in \mathbb{R}^n$ takového, že

$$\inf\{\langle \gamma, y \rangle; y \in K\} > \langle \gamma, x \rangle.$$

Ukažme, že $\alpha = \inf\{\langle \gamma, y \rangle; y \in K\} = 0$. Protože $0 \in K$, dostáváme nerovnost $\alpha \leq 0$. Pokud by existovalo $y \in \mathbb{R}^n$ splňující $\langle \gamma, y \rangle < 0$, pak z limitního přechodu

$$\lim_{t \to \infty} \langle \gamma, ty \rangle = \lim_{t \to \infty} t \langle \gamma, y \rangle = -\infty$$

plyne, že množina $\{\langle \gamma, y \rangle; y \in K\}$ není zdola omezena, což je spor. Tedy $\alpha = 0$, z čehož hned tvrzení plyne.

$\square$

Věta 1.2.4 (Farkasova alternativa)**.** *Nechť A je matice o rozměrech $m \times n$ a nechť $b \in \mathbb{R}^m$. Potom platí právě jedno z následujících tvrzení:*

(i) $\exists x \in \mathbb{R}^n_{+0} : Ax = b$,

(ii) $\exists \gamma \in \mathbb{R}^m : A^T \gamma \geq 0, \langle \gamma, b \rangle < 0$.

Důkaz. Definujme konvexní kužel

$$K = \{\sum_{i=1}^{n} \lambda_i a_i; \lambda_i \geq 0, \ i = 1, ..., n\} \subset \mathbb{R}^m$$

a dokažme, že tento konvexní kužel je uzavřený. Nejprve dokažme uzavřenost množiny

$$M = \{\sum_{i=1}^{n} \lambda_i a_i; \sum_{i=1}^{n} \lambda_i \leq 1, \ \lambda_i \geq 0, \ i = 1, ..., n\}.$$

Zvolme tedy libovolnou posloupnost $\{x_j\}$ bodů v M konvergující k $x \in \mathbb{R}^m$. Pro každé $j \in \mathbb{N}$ a $i = 1, ..., n$ tedy existuje $\lambda_i(x_j) \in [0, 1]$ splňující $x_j = \sum_{i=1}^n \lambda_i(x_j) a_i$. Z kompatknosti $\mathbb{R}^n$ pak můžeme z posloupnosti $\{(\lambda_1(x_j), ..., \lambda_n(x_j)\}$ vybrat podposloupnost konvergující k nějakému $(\lambda_1, ..., \lambda_n)$. Protože platí $x = \sum_{i=1}^n \lambda_i a_i$, $\sum_{i=1}^n \lambda_i \leq 1$ a $\lambda_i \geq 0$ pro $i = 1, ..., n$, je $x \in M$, a tedy M je uzavřená množina.

Zvolme dále libovolnou posloupnost $\{x_j\}$ bodů v K konvergující k $x \in \mathbb{R}^m$. Protože $\{x_j\}$ je cauchyovská, existují $t > 0$ a $j_0 \in \mathbb{N}$ takové, že pro všechna $j \geq j_0$ platí $x_j \in tM$. Z uzavřenosti M plyne $x \in tM$, a tedy $x = \sum_{i=1}^n \lambda_i a_i$, kde pro $i = 1, ..., n$ je $\lambda_i \geq 0$, z čehož plyne, že K je uzavřená množina.

Podle věty o alternativě pro konvexní kužel 1.2.3 platí právě jedno z následujících tvrzení:

(iii) $b \in \overline{K}$,

(iv) $\exists \gamma \in \mathbb{R}^m \; \forall y \in K : \langle \gamma, y \rangle \geq 0, \langle \gamma, b \rangle < 0$.

Pro dokončení důkazu stačí ukázat, že (i) je ekvivalentní s (iii) a (ii) je ekvivalentní s (iv).

To však plyne z ekvivalencí

$$b \in \overline{K} \Leftrightarrow b \in K \Leftrightarrow b = \sum_{i=1}^n \lambda_i a_i, \;\; \lambda_1, ..., \lambda_n > 0 \Leftrightarrow b = Ax \;\; x \in \mathbb{R}^n_{+0}$$

a

$$\begin{aligned} A^T \gamma \geq 0 &\Leftrightarrow \langle a_i, \gamma \rangle \geq 0, \;\; i = 1, ..., n \\ &\Leftrightarrow \langle \lambda_i a_i, \gamma \rangle \geq 0, \lambda_i \geq 0, \;\; i = 1, ..., n \\ &\Leftrightarrow \langle \sum_{i=1}^n \lambda_i a_i, \gamma \rangle \geq 0, \lambda_i \geq 0, \;\; i = 1, ..., n \\ &\Leftrightarrow \langle y, \gamma \rangle \geq 0, \;\; y \in K. \end{aligned}$$

□

Věta 1.2.5 (Stiemkeho alternativa). *Nechť A je matice o rozměrech $m \times n$. Potom platí právě jedno z následujících tvrzení:*

(i) $\exists x \in \mathbb{R}^n_+ : Ax = 0$,

(ii) $\exists y \in \mathbb{R}^m : A^T y > 0$.

Důkaz. Nechť nejdříve platí (i). Pro spor předpokládejme, že existuje $y \in \mathbb{R}^m$ splňující $A^T y > 0$. Protože $x \gg 0$, platí

$$0 < \langle x, A^T y \rangle = \langle Ax, y \rangle = 0,$$

což je spor.

Nechť dále (i) neplatí a najděme řešení pro (ii). Definujme dvě lineární zobrazení:

$$\begin{aligned} T : \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto Ax \end{aligned} \qquad \text{a} \qquad \begin{aligned} T' : \mathbb{R}^m &\to \mathbb{R}^n \\ x &\mapsto A^T x. \end{aligned}$$

Protože neexistuje řešení (i), je $\operatorname{Ker} T \cap \mathbb{R}^n_+ = \emptyset$. Protože obě množiny jsou konvexní, můžeme použít oddělovací větu 1.1.6, která říká, že existuje $\gamma \in \mathbb{R}^n \setminus \{0\}$ splňující

$$\sup\{\langle \gamma, x \rangle; x \in \operatorname{Ker} T\} \leq \inf\{\langle \gamma, y \rangle; y \in \mathbb{R}^n_+\}.$$

Protože $\operatorname{Ker} T$ je lineární prostor, platí pro všechna $x \in \operatorname{Ker} T$ vztah $\langle \gamma, x \rangle = 0$, neboli $\gamma \in (\operatorname{Ker} T)^\perp$. Z lineární algebry víme, že $(\operatorname{Ker} T)^\perp = \operatorname{Im} T'$, a tedy existuje $y \in \mathbb{R}^m$ splňující $A^T y = \gamma$.

Protože pro všechna $y \in \mathbb{R}^n_+$ platí $\langle \gamma, y \rangle \geq 0$, je $\gamma > 0$, a tedy je důkaz dokončen. $\square$

Věta 1.2.6 (Konkávní alternativa). *Nechť $f_1, ..., f_k : K \to \mathbb{R}$ jsou konkávní funkce, kde $K \subset \mathbb{R}^n$ je neprázdná konvexní množina. Definujme $f : K \to \mathbb{R}^k$ předpisem $f(x) = (f_1(x), ..., f_k(x))$. Potom platí právě jedno z následujících tvrzení:*

(i) $\exists x \in K : f(x) \gg 0$,

(ii) $\exists \gamma \in \mathbb{R}^n_{+0}\ \forall x \in K : \langle \gamma, f(x) \rangle \leq 0$.

Důkaz. Pokud platí (i), pak skalární součin v (ii) nabývá kladné hodnoty pro každé $\gamma \in \mathbb{R}^n_{+0}$, čímž dostáváme neplatnost (ii).

Nechť tedy (i) neplatí. Definujme množinu

$$A = \{y \in \mathbb{R}^k; \exists x \in K : y \leq f(x)\}.$$

Množina A je zřejmě neprázdná.

Dále dokažme, že A je konvexní. Nechť $y, z \in A$ splňující $y \leq f(x_1), z \leq f(x_2)$ a nechť $t \in (0,1)$. Protože f_i jsou konkávní, platí

$$ty + (1-t)z \leq tf(x_1) + (1-t)f(x_2) \leq f(tx_1 + (1-t)x_2).$$

Vzhledem k tomu, že K je konvexní, $tx_1 + (1-t)x_2 \in K$, a tedy $ty + (1-t)z \in A$, čímž jsme dokázali konvexitu množiny A.

Z předpokladu plyne, je $A \cap \mathbb{R}^k_+ = \emptyset$. Díky oddělovací větě 1.1.6 existuje $\gamma \in \mathbb{R}^k$ takové, že pro všechny $y \in A$ a $z \in \mathbb{R}^k_+$ platí

$$\langle \gamma, y \rangle \leq \langle \gamma, z \rangle.$$

Protože $\mathbb{R}^k_+ \cup \{0\}$ je kužel, je $\gamma \geq 0$.

Protože $\inf\{\langle \gamma, z \rangle; z \in \mathbb{R}^k_+\} = 0$, máme

$$\langle \gamma, y \rangle \leq 0 \leq \langle \gamma, z \rangle.$$

Vezměme si libovolný bod $x \in K$. Pak $f(x) \in A$ a tedy platí $\langle \gamma, f(x) \rangle \leq 0$, čímž dostáváme požadované tvrzení.

□

Důsledek 1.2.7. *Nechť A je matice o rozměrech $m \times n$. Potom platí právě jedno z následujících tvrzení:*

(i) $\exists x \in \mathbb{R}^n_+ : Ax \gg 0$,

(ii) $\exists y \in \mathbb{R}^m_{+0} : A^T y \leq 0$.

Důkaz. Pro $i = 1, ..., m$ položme $f_i(x) = \langle b_i, x \rangle,\ x \in \mathbb{R}^n_+$, kde b_i jsou řádky matice A. Použitím konkávní alternativy 1.2.6 na $\mathbb{R}^n_{+0}$ a m dostáváme, že buď existuje $x \in \mathbb{R}^n_{+0}$ splňující pro všechna $i = 1, ..., m$

$$0 < f_i(x) = \langle b_i, x \rangle = (Ax)_i$$

nebo existuje $\gamma \in \mathbb{R}^m_{+0}$ takové, že pro všechna $x \in \mathbb{R}^n_{+0}$ platí

$$0 \geq \langle \gamma, Ax \rangle = \langle A^T \gamma, x \rangle.$$

Z prvního případu plyne $Ax \gg 0$, z druhého pak $A^T \gamma \leq 0$ a věta je dokázána.

□

Kapitola 2

Variační principy

2.1 Ekelandův variační princip

Definice 2.1.1. Nechť (X, d) je metrický prostor. Řekneme, že funkce $f : X \to \mathbb{R} \cup \{+\infty\}$ je zdola polospojitá, jestliže pro každou posloupnost $\{x_i\}$ bodů v X konvergující k $x \in X$ platí

$$\liminf_{i\to\infty} f(x_i) \geq f(x).$$

Lemma 2.1.2. *Nechť (X, d) je metrický prostor a nechť $f : X \to \mathbb{R} \cup \{+\infty\}$. Potom f je zdola polospojitá tehdy a jen tehdy, pokud je množina*

$$\{x \in X; f(x) \leq c\}$$

uzavřená pro každé $c \in \mathbb{R}$.

Důkaz. Nechť f je zdola polospojitá funkce. Vezměme si libovolné $c \in \mathbb{R}$ a uvažujme úrovňovou množinu $M = \{x \in X; f(x) \leq c\}$. Pro každou posloupnost $\{x_i\}$ bodů v M konvergující k $x \in X$ pak platí

$$f(x) \leq \liminf_{i\to\infty} f(x_i) \leq c,$$

a tedy $x \in M$ a M je uzavřená.

Nechť pro nějakou $f : X \to \mathbb{R} \cup \{+\infty\}$ platí, že $M = \{x \in X; f(x) \leq c\}$ je uzavřená pro všechna $c \in \mathbb{R}$. Vezměme libovolnou posloupnost $\{x_i\}$ bodů v X konvergující k $x \in X$ a položme $d = \liminf f(x_i)$. Potom existuje taková rostoucí posloupnost indexů $\{i_n\}$ splňující $f(x_{i_n}) \leq d + \frac{1}{n}$. Pro $n < m$ pak platí

$$f(x_{i_m}) \leq d + \frac{1}{m} \leq d + \frac{1}{n}.$$

Z uzavřenosti úrovňových množin ale hned pro každé $n \in \mathbb{N}$ dostáváme $f(x) \leq d + \frac{1}{n}$, a tedy

$$f(x) \leq d = \liminf_{i \to \infty} f(x_i),$$

z čehož plyne, že f je zdola polospojitá.

$\square$

Věta 2.1.3 (Ekelandův variační princip). *Nechť (X, d) je úplný metrický prostor a nechť $f : X \to \mathbb{R} \cup \{+\infty\}$ je zdola polospojitá zdola omezená funkce. Nechť dále $z \in X$ splňuje $f(z) < \inf_{x \in X} f(x) + \varepsilon$ pro nějaké $\varepsilon > 0$. Potom pro každé $\lambda > 0$ existuje $y \in X$, které splňuje následující podmínky:*

(i) $d(z, y) \leq \lambda$,

(ii) $f(y) \leq f(z) - \frac{\varepsilon}{\lambda} d(z, y)$,

(iii) $f(x) \geq f(y) - \frac{\varepsilon}{\lambda} d(x, y),\ x \in X$.

Důkaz. Nechť f splňuje předpoklady věty. Konstruujme induktivně posloupnost $\{z_i\}_{i=0}^{\infty}$ bodů v X a $\{S_i\}_{i=0}^{\infty}$ množin v X. Jako první člen položme $z_0 = z$ a $S_0 = \{x \in X; f(x) + \frac{\varepsilon}{\lambda} d(z, x) \leq f(z)\}$. Jsou-li již prvky $z_0, z_1, ..., z_i$ a množiny $S_0, ..., S_i$ zkonstruovány, vybereme $z_{i+1} \in S_i$ splňující

$$f(z_{i+1}) \leq \frac{1}{2}[f(z_i) + \inf_{x \in S_i} f(x)] \tag{2.1}$$

a volíme

$$S_{i+1} = \{x \in X; f(x) + \frac{\varepsilon}{\lambda} d(z_{i+1}, x) \leq f(z_{i+1})\}.$$

Toto je možné udělat, protože $z_i \in S_i$. Z tohoto dále plyne, že množiny S_i jsou neprázdné pro všechna $i \in \mathbb{N}$.

Protože $f(z)$ je dle předpokladu konečné a pro všechna $i \in \mathbb{N}_0$ platí $z_i \in S_i$, je $\{f(z_i)\}$ nerostoucí posloupnost konečných čísel. Z tohoto plyne, že je konvergentní. Protože $z_{i+1} \in S_i$ podle definice, platí $f(z_{i+1}) + \frac{\varepsilon}{\lambda} d(z_{i+1}, z_i) \leq f(z_i)$, z čehož plyne

$$d(z_{i+1}, z_i) \leq \frac{\lambda}{\varepsilon}[f(z_i) - f(z_{i+1})],\ i \in \mathbb{N}. \tag{2.2}$$

Pro $m > n$ pak platí

$$d(z_n, z_m) \leq \sum_{k=n}^{m-1} d(z_k, z_{k+1}) \leq \sum_{k=n}^{m-1} \frac{\lambda}{\varepsilon}[f(z_k) - f(z_{k+1})] = \frac{\lambda}{\varepsilon}[f(z_n) - f(z_m)]. \tag{2.3}$$

Protože $\{f(z_i)\}$ je cauchyovská, je cauchyovská i posloupnost $\{z_i\}$ a z úplnosti metrického prostoru X dostáváme, že $z_i \to y$ pro nějaké $y \in X$. Dokažme, že toto y je hledaným bodem.

Pokud v (2.3) dosadíme $n = 0$, dostáváme $\frac{\varepsilon}{\lambda} d(z, z_m) \leq f(z) - f(z_m)$. Díky polospojitosti funkce f dostáváme limitním přechodem

$$f(y) + \frac{\varepsilon}{\lambda} d(z, y) \leq \varliminf_{m\to\infty} f(z_m) + \frac{\varepsilon}{\lambda} \lim_{m\to\infty} d(z_m, y) = \varliminf_{m\to\infty} [f(z_m) + \frac{\varepsilon}{\lambda} d(z_m, y)] \leq f(z),$$

a tudíž je podmínka (ii) ověřena.

Tvrzení (i) se pak dostane následující úpravou (ii) :

$$d(z, y) \leq \frac{\lambda}{\varepsilon}[f(z) - f(y)] \leq \frac{\lambda}{\varepsilon}[f(z) - \inf_{x\in X} f(x)] \leq \frac{\lambda}{\varepsilon}\varepsilon = \lambda.$$

Ukážeme nyní, že $\{S_i\}$ je nerostoucí posloupnost množin splňující $\{y\} = \bigcap S_i$. Pro $i \in \mathbb{N}$ a libovolné $x \in S_{i+1}$ totiž využitím (2.2) platí

$$\begin{aligned} f(x) + \frac{\varepsilon}{\lambda} d(x, z_i) &\leq f(x) + \frac{\varepsilon}{\lambda} d(x, z_{i+1}) + \frac{\varepsilon}{\lambda} d(z_{i+1}, z_i) \\ &\leq f(z_{i+1}) + \frac{\varepsilon}{\lambda} d(z_{i+1}, z_i) \leq f(z_i). \end{aligned}$$

Tedy $x \in S_i$ a $S_{i+1} \subseteq S_i$.

Limitním přechodem $m \to \infty$ v (2.3) dostaneme $d(y, z_n) \leq \frac{\lambda}{\varepsilon}[f(z_n) - f(y)]$. Z nerovnosti

$$f(y) + \frac{\varepsilon}{\lambda} d(z_i, y) \leq f(y) + \frac{\varepsilon}{\lambda}\frac{\lambda}{\varepsilon}[f(z_i) - f(y)] = f(z_i)$$

plyne, že $y \in \bigcap_{i=1}^{\infty} S_i$.

Předpokládejme, že $a \in \bigcap_{i=1}^{\infty} S_i$ a dokažme, že pak nutně musí platit $a = y$. Použitím (2.1) dostáváme pro každé $i \in \mathbb{N}$

$$\begin{aligned} \frac{\varepsilon}{\lambda} d(a, z_{i+1}) &\leq f(z_{i+1}) - f(a) \leq f(z_{i+1}) - \inf_{x\in S_i} f(x) \\ &\leq f(z_{i+1}) + f(z_i) - 2f(z_{i+1}) = f(z_i) - f(z_{i+1}). \end{aligned}$$

Z tohoto plyne, že $d(a, y) = 0$, a tedy $a = y$. Protože množiny S_i jsou do sebe vnořené, je toto ekvivalentní se zápisem

$$\forall x \in X\backslash\{y\}\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : x \notin S_n.$$

Máme-li tedy $x \in X\backslash\{y\}$, najdeme $n_0 \in \mathbb{N}$ splňující předchozí. Pak pro $n \geq n_0$ máme $f(x) + \frac{\varepsilon}{\lambda} d(x, z_n) > f(z_n)$ a limitním přechodem $n \to \infty$ dostáváme (iii).

□

Poznámka 2.1.4. Nechť jsou pro $\mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R}, z \in \mathbb{R}^n, \varepsilon > 0$ a $\lambda > 0$ splněny předpoklady Ekelandova variačního principu 2.1.3 a nechť $y \in \mathbb{R}^n$ je nalezený bod. Předpokládejme dále, že f má v bodě y totální diferenciál, který označme $A = f'(y)$. Vezměme libovolný směr $h \in S_{R^n}$. Potom platí

$$\lim_{t \to 0+} \frac{f(y+th) - f(y)}{t} = Ah.$$

Pro libovolné $\delta > 0$ jsme schopni najít $t_0 > 0$ takové, že pro všechny $t \in (0, t_0)$ platí odhad

$$\frac{f(y+th) - f(y)}{t} < Ah + \delta.$$

Vezměme libovolné $t \in (0, t_0)$ a v podmínce (iii) věty 2.1.3 položme $x = y + th$. Potom dostáváme nerovnost $f(y+th) \geq f(y) - \frac{\varepsilon}{\lambda}\|th\|$. Z této nerovnosti ale hned plyne vztah

$$-\frac{\varepsilon}{\lambda} = -\frac{\varepsilon}{\lambda}\frac{\|th\|}{t\|h\|} \leq \frac{f(y+th) - f(y)}{t} < Ah + \delta.$$

Tedy dostáváme

$$-\frac{\varepsilon}{\lambda} < Ah$$

pro každé h o normě 1.

Z linearity A plyne $|Ah| < \frac{\varepsilon}{\lambda}$ a

$$\|f'(y)\| = \sup\{|Ah|; \|h\| = 1\} \leq \frac{\varepsilon}{\lambda}.$$

Tedy pro libovolný bod $y \in \mathbb{R}^n$ splňující $(i), (ii), (iii)$ věty 2.1.3 platí

$$\|f'(y)\| \leq \frac{\varepsilon}{\lambda}.$$

Poznámka 2.1.5. Parametr λ v předchozí větě nám dává jistou možnost volby. Čím menší λ zvolíme, tím bude nalezený bod y blíže původnímu bodu, ale odhad minima či infima funkce se nám o mnoho nezlepší. Pokud bychom naopak zvolili λ velké, dostáváme přesný odhad, ale tato informace bude znehodnocena tím, že nalezený bod bude od původního velmi vzdálen. Často se λ volí rovné 1 nebo $\sqrt{\varepsilon}$.

Věta 2.1.6. *Nechť X je normovaný lineární prostor, $f : X \to \mathbb{R} \cup \{+\infty\}$ zdola omezená zdola polospojitá funkce a $\varepsilon > 0$. Nechť $z \in X$ splňuje $f(z) < \inf_{x \in X} f(x) + \varepsilon$. Potom pro všechna $p \geq 1$ a $\lambda > 0$ existuje $y \in X$ takové, že*

(i) $\|z - y\| \leq \lambda$,

(ii) $f(y) \leq f(z) - \frac{\varepsilon}{\lambda^p}\|z - y\|^p$,

(iii) $f(x) + \frac{\varepsilon}{\lambda^p}\|x - z\|^p \geq f(y) + \frac{\varepsilon}{\lambda^p}\|y - z\|^p, \; x \in X$.

Důkaz. Definujme funkci $g(x) = f(x) + \frac{\varepsilon}{\lambda^p}\|x - z\|^p, x \in X$ a dokažme, že tato funkce nabývá minima na X. Pokud $\|x\| \to \infty$, tak $g(x) \to \infty$. Navíc platí $g(x) \geq \inf_{x \in X} f(x)$, takže g je zdola omezená. Jakožto součet spojité a zdola polospojité funkce je i g zdola polospojitá.

Pro $n \in \mathbb{N}$ definujme posloupnost množin $S_n = \{x \in X; g(x) \leq \inf_{x \in X} g(x) + \frac{1}{n}\}$. Pro všechny $n \in \mathbb{N}$ pak zřejmě platí $S_{n+1} \subseteq S_n$ a z lemmatu 2.1.2 plyne, že S_n jsou uzavřené. Protože pro $\|x\| \to \infty$, platí $g(x) \to \infty$, je $\operatorname{diam}(S_1)$ konečný a tedy $\bigcap_{n=1}^{\infty} S_n \neq \emptyset$.

Funkce g nabývá tedy na X minima, řekněme v bodě $y \in X$, což je přesně tvrzení (iii). Dosazením $x = z$ získáme (ii) a odtud úpravou

$$\|y - z\|^p \leq \frac{f(z) - f(y)}{\varepsilon}\lambda^p \leq \frac{f(z) - \inf_{x \in X} f(x)}{\varepsilon}\lambda^p \leq \frac{\varepsilon}{\varepsilon}\lambda^p = \lambda^p$$

tvrzení (i).

□

Lemma 2.1.7. *Nechť $f : \mathbb{R}^n \to \mathbb{R}$ je zdola omezená funkce, která je na celém definičním oboru spojitě diferencovatelná. Potom existuje posloupnost $\{x_i\}$ bodů v $\mathbb{R}^n$ splňující*

$$f(x_i) \to \inf_{x \in \mathbb{R}^n} f(x) \; \textit{a} \; f'(x_i) \to 0.$$

Důkaz. Najděme libovolnou posloupnost $\{y_i\}$ bodů v $\mathbb{R}^n$ takovou, aby platilo

$$f(y_i) \leq \inf_{x \in \mathbb{R}^n} f(x) + \frac{1}{i}.$$

Pro všechny $i \in \mathbb{N}$ volme $\varepsilon_i = \frac{1}{i}$ a $\lambda_i = 1$. Pro tyto objekty použijme Ekelandův variační princip 2.1.3, čímž dostaneme body $x_i \in \mathbb{R}^n$, které dle (ii) splňují

$$f(x_i) \leq f(y_i) - \frac{\varepsilon_i}{\lambda_i}\|x_i - y_i\| \leq f(y_i) \leq \inf_{x \in \mathbb{R}^n} f(x) + \frac{1}{i} \to \inf_{x \in \mathbb{R}^n} f(x).$$

Dle poznámky o derivaci 2.1.4 pak platí

$$\|f'(x_i)\| \leq \frac{\varepsilon_i}{\lambda_i} = \frac{1}{i} \to 0,$$

čímž je důkaz dokončen.

□

Věta 2.1.8 (Gordan)**.** *Nechť* $a_1, a_2, ..., a_m \in \mathbb{R}^n$*. Potom právě jeden z následujících systémů má řešení:*

$$\sum_{k=1}^{m} \lambda_k a_k = 0,\ \sum_{k=1}^{m} \lambda_k = 1,\ \lambda_k \geq 0,\ k = 1, ..., m, \tag{2.4}$$

$$\langle a_k, x \rangle < 0,\ k = 1, ..., m,\ x \in \mathbb{R}^n. \tag{2.5}$$

Důkaz. Dokažme, že následující podmínky jsou ekvivalentní:

(i) $f(x) = \log(\sum_{k=1}^{m} \exp\langle a_k, x \rangle),\ x \in \mathbb{R}^n$ je omezená zdola,

(ii) systém (2.4) má řešení,

(iii) systém (2.5) nemá řešení.

Dokažme nejdříve sporem $(ii) \Rightarrow (iii)$. Předpokládejme, že $(\lambda_1, ..., \lambda_m)$ je řešení (2.4). Je-li $x \in \mathbb{R}^n$ řešení (2.5), dostávám

$$0 = \langle \sum_{k=1}^{m} \lambda_k a_k, x \rangle = \sum_{k=1}^{m} \lambda_k \langle a_k, x \rangle < 0,$$

což je spor.

Pokud (2.5) nemá řešení, pak pro všechny $x \in \mathbb{R}^n$ existuje index $k_x \in \{1, ..., m\}$ takový, že $\langle a_{k_x}, x \rangle \geq 0$. Z toho ale plyne

$$f(x) = \log(\sum_{k=1}^{m} \exp\langle a_k, x \rangle) \geq \log(\exp\langle a_{k_x}, x \rangle) \geq 0,$$

takže f je zdola omezená.

Zbývá dokázat poslední implikaci. Protože f je spojitě diferencovatelná na $\mathbb{R}^n$ a podle předpokladu zdola omezená, máme podle lemmatu 2.1.7 zaručenu existenci posloupnosti $\{x_i\}$ bodů v $\mathbb{R}^n$ takové, že platí $f(x_i) \to \inf_{x \in \mathbb{R}^n} f(x)$ a $f'(x_i) \to 0$. Pro $k = 1, .., m$ a $i \in \mathbb{N}$ položme

$$\lambda_k^i = \frac{\exp\langle a_k, x_i \rangle}{\sum_{j=1}^{m} \exp\langle a_j, x_i \rangle}.$$

Díky kompaktnosti množiny $S = \{y \in \mathbb{R}^m; \sum_{k=1}^{m} y_k = 1, y_k \geq 0\}$ můžeme bez újmy na obecnosti předpokládat, že $\lambda_k^i \to \lambda_k$ pro všechna $k = 1, ..., m$.

Potom platí

$$f'(x_i) = \frac{1}{\sum_{j=1}^m \exp\langle a_j, x_i\rangle} \sum_{k=1}^m \exp\langle a_k, x_i\rangle a_k$$
$$= \sum_{k=1}^m \frac{\exp\langle a_k, x_i\rangle}{\sum_{j=1}^m \exp\langle a_j, x_i\rangle} a_k = \sum_{k=1}^m \lambda_k^i a_k.$$

Limitním přechodem $i \to \infty$ dostáváme $\sum_{k=1}^m \lambda_k a_k = 0$, což implikuje existenci řešení (2.4). Tím je dokončen důkaz $(i) \Rightarrow (ii)$, a tedy i věty.

$\square$

Poznámka 2.1.9. Podmínka $\sum_{k=1}^m \lambda_k = 1$ v (2.4) je zbytečná, přesto ji uvádím, aby vyniklo, že se jedná o konvexní kombinaci.

Poznámka 2.1.10. Následující větu jsem nalezl v knize [2] na straně 28. Tato věta však neplatí.

Nechť $a_1, a_2, ..., a_m \in \mathbb{R}^n$. Definujme funkci $f : \mathbb{R}^n \to \mathbb{R}$ jako

$$f(x) = \log(\sum_{k=1}^m \exp\langle a_k, x\rangle).$$

Dále pak definujme problémy:

$$\sum_{k=1}^m \lambda_k a_k = 0,\ \sum_{k=1}^m \lambda_k = 1,\ \lambda_k \geq 0,\ k = 1, ..., m, \tag{2.6}$$

$$\langle a_k, x\rangle < 0,\ k = 1, ..., m,\ x \in \mathbb{R}^n_+. \tag{2.7}$$

Potom následující tvrzení jsou ekvivalentní:

(i) $\inf\{f(x), x \geq 0\} > -\infty$,

(ii) systém (2.6) má řešení,

(iii) systém (2.7) nemá řešení.

Volme $m = 1, n = 1, a_1 = 1$. Tím dostáváme $f(x) = x$ a tedy (i) platí. Problém (2.6) se převede na $\lambda_1 = 1, \lambda_1 = 0$, a tedy tento problém nemá řešení. Tím jsme se dostali ke sporu.

Věta 2.1.11 (Stiemke)**.** *Nechť* $a_1, a_2, ..., a_m \in \mathbb{R}^n$. *Definujme funkci*

$$f(x) = \log(\sum_{k=1}^{m} \exp\langle a_k, x\rangle),\ \ x \in \mathbb{R}^n,$$

a problémy

$$\sum_{k=1}^{m} \lambda_k a_k = 0,\ \ \lambda_k > 0,\ \ k = 1, ..., m, \tag{2.8}$$

$$\langle a_k, x\rangle \leq 0,\ \ k = 1, ..., m,\ \ x \in \mathbb{R}^n,\ \ \textit{kde alespoň jedna nerovnost je ostrá.} \tag{2.9}$$

Potom následující podmínky jsou ekvivalentní:

(i) f nabývá minima,

(ii) systém (2.8) *má řešení,*

(iii) systém (2.9) *nemá řešení.*

Důkaz. Dokažme nejdříve sporem $(ii) \Rightarrow (iii)$. Nechť $(\lambda_1, ..., \lambda_m)$ je řešení (2.8) a pro spor předpokládejme, že (2.9) má řešení $x \in \mathbb{R}^n$. Potom ale platí

$$0 = \langle\sum_{k=1}^{m} \lambda_k a_k, x\rangle = \sum_{k=1}^{m} \lambda_k \langle a_k, x\rangle < 0,$$

což je spor.

Důkaz implikace $(iii) \Rightarrow (i)$ rozdělme na dva případy. Prvním bude případ, kdy neexistuje $x \in \mathbb{R}^n \backslash \{0\}$ takové, že pro všechna $k = 1, ..., m$ je $\langle a_k, x\rangle = 0$. To znamená $\mathrm{span}\{a_1, ..., a_m\} = \mathbb{R}^n$. V druhém případě pak takové x existovat bude.

Pro první případ tedy pro každé $x \in S_{\mathbb{R}^n}$ existuje index $k_x \in \{1, ..., m\}$ splňující $\langle a_{k_x}, x\rangle > 0$. Bude tedy existovat okolí $B(x, r_x)$ a kladné číslo δ_x takové, že pro všechny $y \in B(x, r_x)$ platí $\langle a_{k_x}, y\rangle > \delta_x$. Protože $S_{\mathbb{R}^n}$ je kompaktní, existují $x_1, ..., x_j$ splňující

$$\bigcup_{i=1}^{j} B(x_i, r_{x_i}) \supset S_{\mathbb{R}^n}.$$

Položme $\delta = \min\{\delta_{x_1}, ..., \delta_{x_j}\}$. Potom pro všechny $x \in S_{\mathbb{R}^n}$ platí $\langle a_{k_x}, x\rangle > \delta$, a tedy pro $t > 0$ platí

$$f(tx) = \log(\sum_{k=1}^{m} \exp\langle a_k, tx\rangle) \geq \langle a_{k_x}, tx\rangle > t\delta.$$

Z tohoto plyne, že $\lim_{\|x\|\to\infty} f(x) = \infty$, a tedy f nabývá na $\mathbb{R}^n$ minima.

Pro druhý případ položme

$$A = \operatorname{span}\{a_1, ..., a_m\} \quad \text{a} \quad B = A^\perp.$$

Z definice ortogonálního doplňku můžeme pro každé $x \in \mathbb{R}^n$ najít $x_A \in A$ a $x_B \in B$ takové, že platí $x = x_A + x_B$. Potom

$$\begin{aligned} f(x) = \log(\sum_{k=1}^m \exp\langle a_k, x\rangle) &= \log(\sum_{k=1}^m \exp\langle a_k, x_A + x_B\rangle) \\ = \log(\sum_{k=1}^m \exp\langle a_k, x_A\rangle \exp\langle a_k, x_B\rangle) &= \log(\sum_{k=1}^m \exp\langle a_k, x_A\rangle) = f(x_A). \end{aligned}$$

Tedy $f(A) = f(\mathbb{R}^n)$, čímž jsme problém převedli na první případ, kde už máme existenci minima zaručenou.

Zbývá dokázat poslední implikace $(i) \Rightarrow (ii)$. Protože f je spojitě diferencovatelná a nabývá minima v nějakém bodě $z \in \mathbb{R}^n$, musí platit $f'(z) = 0$. Pro $k = 1, ..., m$ definujme

$$\lambda_k = \frac{\exp\langle a_k, z\rangle}{\sum_{j=1}^m \exp\langle a_j, z\rangle} > 0.$$

a dokažme, že tyto λ_k jsou řešením (2.9):

$$\begin{aligned} 0 = f'(z) &= \frac{1}{\sum_{j=1}^m \exp\langle a_j, z\rangle} \sum_{k=1}^m \exp\langle a_k, z\rangle a_k \\ &= \sum_{k=1}^m \frac{\exp\langle a_k, z\rangle}{\sum_{j=1}^m \exp\langle a_j, z\rangle} a_k = \sum_{k=1}^m \lambda_k a_k. \end{aligned}$$

□

2.2 Borwein-Preissův variační princip

Definice 2.2.1. Nechť (X, d) je metrický prostor. Řekneme, že funkce $\rho : X \times X \to [0, \infty)$ je měrka, pokud jsou splněny následující podmínky:

(i) ρ je spojitá,

(ii) $\rho(x, x) = 0,\ x \in X$,

(iii) $\forall \varepsilon > 0\ \exists \delta > 0\ \forall y, z \in X : \rho(y, z) \leq \delta \Rightarrow d(y, z) \leq \varepsilon$,

(iv) $\rho(x, y) = \rho(y, x),\ x, y \in X$.

Poznámka 2.2.2. Tato kapitola je převzaná z [2]. V definici 2.2.1 se v této knize nepožaduje podmínka (*iv*). Protože však budeme tuto vlastnost potřebovat v další sekci, zařadil jsem ji do definice.

Věta 2.2.3 (Borwein-Preiss)**.** *Nechť (X, d) je úplný metrický prostor a nechť $f : X \to \mathbb{R} \cup \{+\infty\}$ je zdola omezená zdola polospojitá funkce. Nechť ρ je měrka a nechť $\{\delta_i\}_{i=0}^{\infty}$ je posloupnost kladných čísel. Nechť existují $\varepsilon > 0$ a $z \in X$ splňující*

$$f(z) < \inf_{x \in X} f(x) + \varepsilon.$$

Potom existuje $y \in X$ a posloupnost $\{x_i\}_{i=0}^{\infty}$ bodů v X s následujícími vlastnostmi:

(i) $\rho(z, y) \leq \frac{\varepsilon}{\delta_0}, \rho(x_i, y) \leq \frac{\varepsilon}{2^i \delta_0},\ i \in \mathbb{N}_0$,

(ii) $f(y) + \sum_{i=0}^{\infty} \delta_i \rho(y, x_i) \leq f(z)$,

(iii) $f(x) + \sum_{i=0}^{\infty} \delta_i \rho(x, x_i) \geq f(y) + \sum_{i=0}^{\infty} \delta_i \rho(y, x_i)$.

Důkaz. Pro $n \in \mathbb{N}_0$ induktivně definujme body $x_n \in X$ a množiny $S_n \subset X$. Jako první členy položme $x_0 = z$ a $S_0 = \{x \in X; f(x) + \delta_0 \rho(x, x_0) \leq f(x_0)\}$. Protože $x_0 \in S_0$, je S_0 neprázdná.

Pokud již máme definovány body $x_0, ..., x_n$ a množiny $S_0, ..., S_n$, vyberme bod $x_{n+1} \in S_n$ splňující

$$f(x_{n+1}) + \sum_{k=0}^{n} \delta_k \rho(x_{n+1}, x_k) \leq \inf_{x \in S_n} [f(x) + \sum_{k=0}^{n} \delta_k \rho(x, x_k)] + \frac{\varepsilon \delta_{n+1}}{2^{n+1} \delta_0}.$$

Toto je možné udělat, protože $f(x) + \sum_{k=0}^{n} \delta_k \rho(x, x_k)$ je zdola omezená funkce. Dále pak definujme množinu

$$S_{n+1} = \{x \in S_n; f(x) + \sum_{k=0}^{n+1} \delta_k \rho(x, x_k) \leq f(x_{n+1}) + \sum_{k=0}^{n} \delta_k \rho(x_{n+1}, x_k)\},$$

čímž je induktivní konstrukce dokončena.

Zkoumejme pro $n \in \mathbb{N}_0$ vlastnosti množin S_n. Protože $f(z) < \infty$ dle předpokladu a $x_n \in S_n$, je $f(x_n)$ konečné číslo. Zřejmě jsou S_n neprázdné. Z lemmatu 2.1.2 plyne, že S_n jsou uzavřené. Pro libovolné $x \in S_n$ pak platí

$$\begin{aligned}\delta_n \rho(x, x_n) &\leq [f(x_n) + \sum_{k=0}^{n-1} \delta_k \rho(x_n, x_k)] - [f(x) + \sum_{k=0}^{n-1} \delta_k \rho(x, x_k)] \\ &\leq [f(x_n) + \sum_{k=0}^{n-1} \delta_k \rho(x_n, x_k)] - \inf_{x \in S_{n-1}} [f(x) + \sum_{k=0}^{n-1} \delta_k \rho(x, x_k)] \leq \frac{\varepsilon \delta_n}{2^n \delta_0},\end{aligned}$$

tedy

$$\rho(x, x_n) \leq \frac{\varepsilon}{2^n \delta_0}. \tag{2.10}$$

Ukažme, že $\operatorname{diam}(S_n) \to 0$. Zvolme si $\varepsilon' > 0$ libovolné a z definice měrky najděme odpovídající δ'. Zvolme $n_0 \in \mathbb{N}$ splňující $\frac{\varepsilon}{2^{n_0} \delta_0} < \delta'$. Potom pro $n \geq n_0$ a $x \in S_n$ máme z (2.10) nerovnost $d(x, x_n) \leq \varepsilon'$. Tedy $\operatorname{diam}(S_n) \leq 2\varepsilon'$ pro $n \geq n_0$.

Protože S_n jsou neprázdné, uzavřené a jejich diametr konverguje k nule, existuje právě jedno $y \in X$ splňující $y \in \bigcap_{n=0}^{\infty} S_n$. Z (2.10) a $z \in S_0$ dostáváme $x_i \to y$ a (i).

Protože pro libovolné $n \in \mathbb{N}$ platí $x_n \in S_{n-1}$, dostáváme nerovnosti

$$\begin{aligned}f(z) = f(x_0) \geq f(x_1) + \delta_0 \rho(x_1, x_0) \geq f(x_2) + \sum_{k=0}^{1} \delta_k \rho(x_1, x_k) \\ \geq \dots \geq f(x_n) + \sum_{k=0}^{n-1} \delta_k \rho(x_n, x_k).\end{aligned}$$

Limitním přechodem $n \to \infty$ potom dostáváme (ii).

Pokud si vezmeme libovolné $x \in X \backslash \{y\}$, pak $x \notin \bigcap_{n=0}^{\infty} S_n$, a tedy existuje $n_0 \in \mathbb{N}$ takové, že pro všechny $n \geq n_0$ platí

$$f(x) + \sum_{k=0}^{\infty} \delta_k \rho(x, x_k) \geq f(x) + \sum_{k=0}^{n} \delta_k \rho(x, x_k) > f(x_n) + \sum_{k=0}^{n-1} \delta_k \rho(x_n, x_k).$$

Limitním přechodem

$$f(x) + \sum_{k=0}^{\infty} \delta_k \rho(x, x_k) \geq \liminf_{n \to \infty} [f(x_n) + \sum_{k=0}^{n-1} \delta_k \rho(x_n, x_k)] \geq f(y) + \sum_{k=0}^{\infty} \delta_k \rho(y, x_k)$$

dostáváme (iii).

$\square$

Poznámka 2.2.4. Dokážeme teď z Borwein-Preissova variačního principu o trochu slabší variantu Ekelandova variačního principu. Nechť (X,d) je úplný metrický prostor a nechť $f : X \to \mathbb{R} \cup \{+\infty\}$ je zdola polospojitá zdola omezená funkce. Nechť dále $z \in X$ splňuje $f(z) < \inf_{x\in X} f(x) + \varepsilon$ pro nějaké $\varepsilon > 0$.

Definujme $\rho : X \times X \to [0,\infty)$ jako $\rho(x,y) = \frac{\varepsilon}{\lambda} d(x,y)$ a zvolme δ_i kladná splňující $\sum_{i=0}^{\infty} \delta_i = 1$. Zřejmě ρ je měrka, použijme tedy na objekty $(X,d), \rho, f, z, \varepsilon$ větu 2.2.3. Dostaneme posloupnost $\{x_i\}$ bodů v X konvergující k nějakému $y \in X$, pro které bude platit:

(i) $\frac{\varepsilon}{\lambda} d(z,y) \leq \frac{\varepsilon}{\delta_0}$,

(ii) $f(y) + \sum_{i=0}^{\infty} \delta_i \frac{\varepsilon}{\lambda} d(y,x_i) \leq f(z)$,

(iii) $f(x) + \sum_{i=0}^{\infty} \delta_i \frac{\varepsilon}{\lambda} d(x,x_i) \geq f(y) + \sum_{i=0}^{\infty} \delta_i \frac{\varepsilon}{\lambda} d(y,x_i)$.

Jednotlivé tvrzení upravme, v (iii) použijme trojúhelníkovou nerovnost. Potom dostaneme:

(i) $d(z,y) \leq \frac{1}{\delta_0}\lambda$,

(ii) $f(y) + \delta_0 \frac{\varepsilon}{\lambda} d(y,z) \leq f(z)$,

(iii) $f(x) \geq f(y) - \frac{\varepsilon}{\lambda} d(y,z)$.

Protože můžeme zvolit libovolné $\delta_0 \in (0,1)$, dostáváme o něco slabší variantu Ekelandova variačního principu.

2.3 Lindenstrauss-Preiss-Tišerův variační princip

Definice 2.3.1. Nechť (X,d) je metrický prostor a nechť d_0 je spojitá pseudometrika na X. Řekneme, že X je (d,d_0)-úplný, pokud existují funkce

$$\delta_j : X^{j+1} \to (0,\infty),\ j \in \mathbb{N}_0$$

takové, že každá d-cauchyovská posloupnost $\{x_j\}_{j=0}^{\infty}$ v X splňující

$$d_0(x_j, x_{j+1}) \leq \delta_j(x_0, ..., x_j),\ j \in \mathbb{N}_0$$

Je konvergentní v (X,d).

Definice 2.3.2. Nechť (X, d) je metrický prostor a nechť d_0 je spojitá pseudometrika na X. Řekneme, že $f : X \to \mathbb{R}$ je (d, d_0)-zdola polospojitá, pokud existují funkce

$$\delta_j : X^{j+1} \to (0, \infty),\ j \in \mathbb{N}_0$$

takové, že pro každou posloupnost bodů $\{x_j\}_{j=0}^{\infty}$ v X konvergující k nějakému $x \in X$ splňující

$$d_0(x_j, x_{j+1}) \le \delta_j(x_0, ..., x_j),\ j \in \mathbb{N}_0$$

platí, že

$$f(x) \le \liminf_{j\to\infty} f(x_j).$$

Poznámka 2.3.3. Pokud nebude uvedeno jinak, budeme v dalším průběhu kapitoly všechny metrické vlastnosti vztahovat k metrice d.

Lemma 2.3.4. *Nechť (X, d) je metrický prostor a nechť d_0 je spojitá pseudometrika na X. Dále nechť $f : X \to \mathbb{R}$ splňuje, že pro každé $r \in \mathbb{R}$ je množina $\{x \in X, f(x) \le r\}$ typu G_δ v (X, d_0). Potom f je (d, d_0)-zdola polospojitá.*

Důkaz. Množina je typu G_δ, pokud ji můžeme zapsat jako spočetný průnik otevřených množin. Pro každé $q \in \mathbb{Q}$ můžeme tedy napsat

$$\{x \in X; f(x) \le q\} = \bigcap_{i=1}^{\infty} H_i^q,$$

kde H_i^g jsou d_0-otevřené množiny. Protože množin $H_i^q \subset X$ je spočetně mnoho, můžeme je očíslovat. Udělejme tak a označme je $G_0, G_1, ...$

Definujme funkci $T : X \times \{G_i, i \in \mathbb{N}_0\} \to (0, \infty]$ následovně:

$$T(x, G_i) = \begin{cases} \frac{1}{4}\sup\{r > 0; B_{d_0}(x, r) \subset G_i\}, & x \in G_i\ , \\ \infty, & x \notin G_i\ . \end{cases}$$

Protože G_i jsou otevřené pro všechny $i \in \mathbb{N}_0$, zobrazuje T skutečně do $(0, \infty]$.

Zkonstruujme induktivně funkce $\delta_j : X^{j+1} \to (0, \infty),\ j \in \mathbb{N}_0$, které budou splňovat následující podmínky:

(i) $\delta_j(x_0, ..., x_j) \le \frac{1}{2}\delta_{j-1}(x_0, ..., x_{j-1}),\ j \in \mathbb{N}$,

(ii) $B_{d_0}(x_j, 2\delta_j(x_0, ..., x_j)) \subset G_i,\ 0 \le i < j,\ x_j \in G_i$.

Jako první člen položme $\delta_0(x) = 1$ pro všechny $x \in X$. Pokud již máme definováno $\delta_0, ..., \delta_{j-1}$, položme

$$\delta_j(x_0, ..., x_j) = \min\{\frac{1}{2}\delta_{j-1}(x_0, ..., x_{j-1}), T(x_j, G_0), ..., T(x_j, G_{j-1})\},\ x_0, ..., x_j \in X,$$

čímž dostaneme funkce požadovaných vlastností.

Zvolme posloupnost $\{x_j\}$ bodů v X konvergující k nějakému $x_\infty \in X$ splňující

$$d_0(x_j, x_{j+1}) \le \delta_j(x_0, ..., x_j),\ j \in \mathbb{N}_0.$$

Dále zvolme libovolné $q \in \mathbb{Q}$ splňující $q < f(x_\infty)$ a najděme G_i, která splňuje

$$\{x \in X, f(x) \le q\} \subset G_i,\ x_\infty \notin G_i.$$

Dokažme, že pro $j > i$ platí $f(x_j) > q$. Pokud by existovalo $j > i$, pro které by platilo $f(x_j) \le q$, pak $x_j \in G_i$ a dle volby δ_j dostáváme

$$B_{d_0}(x_j, 2\delta_j(x_0, ..., x_j)) \subset G_i.$$

Současně ale z volby δ_j pro $k > j$ plyne

$$\begin{aligned} d_0(x_j, x_k) &\le \sum_{l=j}^{k-1} d_0(x_l, x_{l+1}) \le \sum_{l=j}^{k-1} \delta_l(x_0, ..., x_l) \\ &\le \sum_{l=j}^{k-1} 2^{j-l}\delta_j(x_0, ..., x_j) \le 2\delta_j(x_0, ..., x_j), \end{aligned}$$

tedy

$$d_0(x_j, x_\infty) \le 2\delta_j(x_0, ..., x_j),$$

z čehož hned dostáváme $x_\infty \in B_{d_0}(x_j, 2\delta_j(x_0, ..., x_j)) \subset G_i$, což je spor.

Ukázali jsme, že $\liminf f(x_i) \ge q$ pro všechny $q \in \mathbb{Q}$ splňující $q < f(x_\infty)$. Tedy

$$f(x_\infty) \le \liminf_{j\to\infty} f(x_j),$$

čímž je věta dokázána.

□

Lemma 2.3.5. *Nechť (X, d) je úplný metrický prostor, nechť d_0 je spojitá pseudometrika na X a nechť S je G_δ podmnožina (X, d_0). Potom S je (d, d_0)-úplná.*

Důkaz. Nechť f je charakteristická funkce množiny $X \backslash S$. Pro $r \in \mathbb{R}$ je

$$\{x \in X, f(x) \leq r\}$$

je prázdná množina, S nebo X, je tedy G_δ podmnožina (X, d_0). Tím jsou splněny předpoklady lemmatu 2.3.4, tedy f je (d, d_0)-zdola polospojitá. Tím dostáváme pro $j \in \mathbb{N}_0$ existenci funkcí $\delta_j : X^{j+1} \to (0, \infty)$, které splňují podmínky z definice 2.3.2.

Mějme cauchyovskou posloupnost $\{x_j\}$ bodů v S splňující pro $j \in \mathbb{N}_0$

$$d_0(x_j, x_{j+1}) \leq \delta_j(x_0, ..., x_j).$$

Z úplnosti X dostáváme $x_j \to x$ pro nějaké $x \in X$. Protože f je (d, d_0)-zdola polospojitá, dostáváme

$$f(x) \leq \liminf_{j \to \infty} f(x_j) = 0.$$

Tedy $f(x) = 0$, neboli $x \in S$. Tím jsme ovšem dokázali, že S je (d, d_0)-úplná.

□

Věta 2.3.6 (Lindenstrauss-Preiss-Tišerův variační princip). *Nechť (X, d) je metrický prostor a nechť pro $j \in \mathbb{N}_0$ splňují funkce*

$$F_j : X \times X \to [0, \infty]$$

následující podmínky:

(a) F_j je d-zdola polospojitá v druhé proměnné,

(b) $F_j(x, x) = 0$ pro všechny $x \in X$,

(c) existuje nerostoucí posloupnost $\{r_j\}$ splňující $\lim r_j = 0$ a $\inf_{d(x,y) > r_j} F_j(x, y) > 0$.

Dále nechť d_0 je spojitá pseudometrika na X, $S \subset X$ je (d, d_0)-úplná a $x_0 \in S$. Nechť $f : X \to \mathbb{R}$ je (d, d_0)-zdola polospojitá zdola omezená funkce. Nechť $\{\varepsilon_j\}_{j=0}^{\infty}$ je posloupnost kladných čísel, pro kterou je splněno

$$f(x_0) \leq \varepsilon_0 + \inf_{x \in S} f(x).$$

Potom existuje posloupnost $\{x_j\}_{j=1}^{\infty}$ bodů v S, která v metrice d konverguje k nějakému $x_\infty \in S$ a d_0-spojitá funkce $\varphi : X \to \mathbb{R}$ taková, že funkce

$$h(x) = f(x) + \varphi(x) + \sum_{j=0}^{\infty} F_j(x_j, x)$$

nabývá minima na S v bodě x_∞.

Dále pro $0 \leq j < k+1 \leq \infty$ platí

$$\begin{aligned} F_j(x_j, x_k) &\leq \varepsilon_j, \\ h(x_\infty) &\leq \varepsilon_j + \inf_{x \in S}[f(x) + \varphi(x) + \sum_{i=0}^{j-1} F_i(x_i, x)]. \end{aligned}$$

Důkaz. Bez újmy na obecnosti předpokládejme, že $\sum_{j=1}^{\infty} \varepsilon_j < \infty$ a

$$\inf_{d(x,y)>r_j} F_j(x,y) > \varepsilon_j, \ j \in \mathbb{N}.$$

Dále předpokládejme, že funkce $\delta_j : M^{j+1} \to (0, \infty), \ j \in \mathbb{N}_0$ splňují jak podmínky 2.3.1, tak i 2.3.2.

Pro $j \in \mathbb{N}_0$ induktivně zkonstruujme d_0-zdola polospojité funkce $\varphi_j : X \to [0, \infty)$ a zdola omezené funkce $h_j : X \to (-\infty, \infty]$. Pro $j \in \mathbb{N}$ induktivně definujme body $x_j \in S$. Pro $j = 0$ položme

$$\begin{aligned} h_0(x) &= f(x), \\ \varphi_0(x) &= 2\varepsilon_0 \min\{1, \frac{d_0(x, x_0)}{\delta_0(x_0)}\}, \ x \in X. \end{aligned}$$

Nechť $j \in \mathbb{N}$ a nechť jsou x_i, φ_i a h_i definovány pro všechna $i = 0, ..., j-1$. Potom definujme

$$h_j(x) = h_{j-1}(x) + \varphi_{j-1}(x) + F_{j-1}(x_{j-1}, x), \ x \in X.$$

Následně zvolme takové $x_j \in S$, které bude splňovat

$$h_j(x_j) \leq \min\{h_{j-1}(x_{j-1}), \varepsilon_j + \inf_{x \in S} h_j(x)\}.$$

Toto je možné udělat, protože $h_j(x_{j-1}) = h_{j-1}(x_{j-1})$. Nakonec položme

$$\varphi_j(x) = 2\varepsilon_j \min\{1, \frac{d_0(x, x_j)}{\delta_j(x_0, ..., x_j)}\},$$

čímž je induktivní konstrukce dokončena.

Potom pro tyto objekty platí následující:

(i) $\varphi_j(x_j) = 0, \ j \in \mathbb{N}_0$,

(ii) $h_j(x_j) \leq \varepsilon_j + \inf_{x \in S} h_j(x), \ j \in \mathbb{N}_0$,

(iii) $h_j(x_j) \leq h_{j-1}(x_{j-1}),\ j \in \mathbb{N}$,

(iv) $h_j(x) \geq h_{j-1}(x),\ j \in \mathbb{N},\ x \in S$,

(v) $\varphi_i(x_j) + F_i(x_i, x_j) \leq \varepsilon_i,\ j \geq i,\ j, i \in \mathbb{N}_0$.

Vlastnosti $(i) - (iv)$ jsou zřejmé, zbývá tedy dokázat ještě vlastnost (v). Pokud platí $i = j$, pak je levá strana nulová. Pokud $j > i \geq 0$, dostáváme

$$\begin{aligned} h_i(x_i) \geq h_j(x_j) \geq h_{i+1}(x_j) &= h_i(x_j) + \varphi_i(x_j) + F_i(x_i, x_j) \\ &\geq h_i(x_i) + \varphi_i(x_j) + F_i(x_i, x_j) - \varepsilon_i, \end{aligned}$$

z čehož hned plyne vlastnost (v).

Dále platí

$$\begin{aligned} h_j(x) &= h_{j-1}(x) + \varphi_{j-1}(x) + F_{j-1}(x_{j-1}, x) = ... = \\ &= h_0(x) + \sum_{i=0}^{j-1} [\varphi_i(x) + F_i(x_i, x)] = f(x) + \sum_{i=0}^{j-1} [\varphi_i(x) + F_i(x_i, x)]. \end{aligned}$$

Pro libovolné $\varepsilon > 0$ najdeme r_i, které splňuje $r_i < \varepsilon$. Z (v) plyne $F_i(x_i, x_j) \leq \varepsilon_i$ pro $j \geq i$, z čehož použitím (c) dostáváme

$$d(x_i, x_j) \leq r_i < \varepsilon,$$

a tedy $\{x_i\}$ je d-cauchyovská. Dále pak pro $j \geq i$ platí

$$d_0(x_i, x_j) \leq \delta_i(x_0, ..., x_i).$$

Pokud by toto nebyla pravda, platilo by $\varphi_i(x_j) = 2\varepsilon_i$, což je spor s (v).

Z (d, d_0)-úplnosti a (d, d_0)-zdola polospojitosti funkce f plyne, že $\{x_j\}$ konverguje k nějakému $x_\infty \in S$ a pro tento bod platí

$$f(x_\infty) \leq \liminf_{j \to \infty} f(x_j).$$

Položme $\varphi(x) = \sum_{j=0}^{\infty} \varphi_j(x),\ x \in X$. Potom pro všechna $x \in X$ platí

$$\varphi(x) = \sum_{j=0}^{\infty} \varphi_j(x) = \sum_{j=0}^{\infty} 2\varepsilon_j \min\{1, \frac{d_0(x, x_j)}{\delta_j(x_0, ..., x_j)}\} \leq \sum_{j=0}^{\infty} 2\varepsilon_j < \infty,$$

tedy φ je korektně definovaná a d-spojitá funkce. Položme

$$h(x) = f(x) + \varphi(x) + \sum_{j=0}^{\infty} F_j(x_j, x).$$

Protože F_j jsou d-zdola polospojité v druhé proměnné, pro libovolné pevné $k \in \mathbb{N}$ platí:

$$h_0(x_0) \geq h_1(x_1) \geq \lim_{j\to\infty} h_j(x_j) \geq \liminf_{j\to\infty} [f(x_j) + \varphi(x_j) + \sum_{i=0}^{k} F_i(x_i, x_j)] \geq$$

$$\geq f(x_\infty) + \varphi(x_\infty) + \sum_{i=0}^{k} F_i(x_i, x_\infty).$$

Limitním přechodem $k \to \infty$ dostáváme pro všechna $j \in \mathbb{N}_0$ nerovnost $h_j(x_j) \geq h(x_\infty)$.

Ze (v) ihned pro $j \geq i \geq 0$ plyne $F_i(x_i, x_j) \leq \varepsilon_i$. Z (ii) potom dostáváme

$$h(x_\infty) \leq h_j(x_j) \leq \varepsilon_j + \inf_{x\in S} h_j(x) = \varepsilon_j + \inf_{x\in S} \{f(x) + \sum_{i=0}^{j-1} [\varphi_i(x) + F_i(x_i, x)]\}$$

$$\leq \varepsilon_j + \inf_{x\in S} [f(x) + \varphi(x) + \sum_{i=0}^{j-1} F_i(x_i, x)]$$

a

$$h(x_\infty) \leq \varepsilon_j + \inf_{x\in S} [f(x) + \varphi(x) + \sum_{i=0}^{j-1} F_i(x_i, x)] \leq \varepsilon_j + \inf_{x\in S} h(x),$$

a tedy h nabývá na S minima v x_∞.

□

Poznámka 2.3.7. Pokud jsou splněny předpoklady věty 2.3.6, najdeme nějakou φ, pomocí které je definována h. Bez újmy na obecnosti můžeme předpokládat $\varphi(x_\infty) = 0$. Pro $x \in X$ definujme $\varphi'(x) = \max\{0, \varphi(x)\}$ a

$$h'(x) = f(x) + \varphi'(x) + \sum_{j=0}^{\infty} F_j(x_j, x).$$

Potom pro všechny $x \in S$ platí

$$h'(x_\infty) = h(x_\infty) \leq h(x) = f(x) + \varphi(x) + \sum_{j=0}^{\infty} F_j(x_j, x) \leq$$

$$\leq f(x) + \varphi'(x) + \sum_{j=0}^{\infty} F_j(x_j, x) = h'(x),$$

a tedy h' nabývá na S minimum v x_∞.

Můžeme tedy ve větě 2.3.6 požadovat existenci φ nezáporné s vlastností $\varphi(x_\infty) = 0$.

Poznámka 2.3.8. Dále se zabývejme případem, kdy $d_0 = 0$, což je zřejmě pseudometrika. Podmínka na S a f se změní na d-úplnost a d-zdola polospojitost. Z definice d_0-spojitosti dostaneme, že φ je konstatní funkce. Podle poznámky 2.3.7 můžeme tedy volit φ nulovou.

Poznámka 2.3.9. Předpokládejme, že v předpokladech věty 2.3.6 platí v podmínce $f(x_0) \leq \varepsilon_0 + \inf_{x\in S} f(x)$ ostrá nerovnost. Omezme se ve větě 2.2.3 na případ, kdy f nabývá pouze konečných hodnot. Dokažme poté, že v tomto případě je věta 2.3.6 zobecněním věty 2.2.3.

Nechť jsou splněny předpoklady Borwein-Preissova variačního principu. Mějme tedy úplný metrický prostor (X, d), zdola omezenou zdola polospojitou funkci $f : X \to \mathbb{R}$ posloupnost $\{\delta_i\}$ kladných čísel, bod $x \in X$, pro který platí $f(x) < \inf_{x\in X} f(x) + \varepsilon$, kde $\varepsilon > 0$ a $\rho : X \times X \to [0, \infty)$ měrku.

V předchozí větě zvolme postupně

$$S = X,\ d_0 = 0,\ F_j(x, y) = \delta_j \rho(y, x),\ j \in \mathbb{N}_0,$$

$$\varepsilon_0 = f(z) - \inf_{x\in X} f(x),\ \varepsilon_j = \frac{\delta_j}{\delta_0 2^j}\varepsilon,\ j \in \mathbb{N}.$$

Poté jsou splněny předpoklady věty 2.3.6.

Protože jsme zvolili $d_0 = 0$, můžeme zvolit $\varphi = 0$. Potom funkce

$$h(x) = f(x) + \sum_{j=0}^{\infty} \delta_j \rho(x, x_j)$$

nabývá minima v $x_\infty = y$, což je přesně (iii) v Borwein-Preissově variačním principu.

Z nerovnosti $F_j(x_j, x_k) \leq \varepsilon_j$ pro všechny $j \in \mathbb{N}_0$ limitním přechodem $k \to \infty$ plyne $\rho(y, x_j) \leq \frac{\varepsilon}{\delta_0 2^j}$. Pokud zvolíme $j = 0$ dostaneme $\rho(y, z) \leq \frac{\varepsilon}{\delta_0}$. Díky symetrii ρ dostáváme (i).

Z nerovnosti

$$h(y) \leq \varepsilon_i + \inf_{x\in X} [f(x) + \sum_{j=0}^{i-1} F_i(x_i, x)]$$

dostaneme volbou $i = 0$

$$f(y) + \sum_{j=0}^{\infty} \delta_j \rho(y, x_j) = h(y) \leq \varepsilon_0 + \inf_{x\in X} f(x) = f(z),$$

což je přesně (ii).

2.4 Vztah s úplností

Věta 2.4.1 (Ekelandův variační princip a úplnost)**.** *Nechť (X, d) je metrický prostor. Potom X je úplný právě tehdy, pokud pro každou zdola omezenou zdola polospojitou funkci $f : X \to \mathbb{R}$ platí Ekelandův variační princip, tedy pro libovolné $\varepsilon > 0$ existuje $y \in X$ splňující*

(i) $f(y) \leq \inf_{x \in X} f(x) + \varepsilon$,

(ii) $f(x) > f(y) - \varepsilon d(y, x), \ x \in X \setminus \{y\}$.

Důkaz. Nechť (X, d) je úplný metrický prostor a nechť f je zdola omezená zdola polospojitá funkce. Volme $\lambda = 1$ a najděme $z \in X$ splňující $f(z) \leq \inf_{x \in X} f(x) + \varepsilon$. Použitím věty 2.1.3, dostáváme, že existuje $y \in X$ s požadovanými vlastnostmi.

Pro druhou implikaci uvažujme cauchyovskou posloupnost $\{x_i\}$ bodů v X. Pro $x \in X$ definujme

$$f(x) = \lim_{i \to \infty} d(x_i, x).$$

Protože pro všechna $m, n \in \mathbb{N}$ platí $|d(x_m, x) - d(x_n, x)| \leq d(x_m, x_n)$, je f korektně definována.

Ověřme předpoklady Ekelandova variačního principu 2.1.3. Pro libovolné body $a, b \in X$ pak platí

$$\begin{aligned} |f(a) - f(b)| &= |\lim_{i \to \infty} d(x_i, a) - \lim_{i \to \infty} d(x_i, b)| = |\lim_{i \to \infty} [d(x_i, a) - d(x_i, b)]| \\ &\leq \lim_{i \to \infty} |d(x_i, a) - d(x_i, b)| \leq \lim_{i \to \infty} |d(a, b)| = |d(a, b)|. \end{aligned}$$

Tedy f je $1-$lipschitzovská, což znamená, že je i spojitá. Dále je f je zdola omezená nulou a pro posloupnost $\{x_i\}$ platí $f(x_i) \to 0$, takže $\inf_{x \in X} f(x) = 0$.

Z předpokladů věty víme, že

$$\forall \varepsilon > 0 \ \exists y \in X, f(y) \leq \varepsilon \ \forall x \in X : f(y) \leq f(x) + \varepsilon d(x, y).$$

Zvolme si $\varepsilon \in (0, 1)$ a $x = x_i$. Najdeme $y \in X$ s požadovanou vlastností. Protože $f(y) \leq f(x_i) + \varepsilon d(x_i, y)$, limitním přechodem $i \to \infty$ dostáváme

$$f(y) \leq 0 + \varepsilon f(y),$$

z čehož plyne $0 = f(y) = \lim_{i \to \infty} d(x_i, y)$, takže $x_i \to y$ a X je úplný.

□

Definice 2.4.2. Nechť (X, d) je metrický prostor a $\phi : X \to X$. Řekneme, že ϕ je kontrakce, pokud existuje $k \in (0, 1)$ takové, že pro všechna $x, y \in X$ platí

$$d(\phi(x), \phi(y)) \leq kd(x, y).$$

Poznámka 2.4.3. Každá kontrakce je spojitá funkce, neboť je k-lipschitzovské zobrazení.

Věta 2.4.4 (Banachova věta o kontrakci)**.** *Nechť (X, d) je úplný metrický prostor a nechť $\phi : X \to X$ je kontrakce. Potom ϕ má právě jeden pevný bod.*

Důkaz. Definujme $f : X \to \mathbb{R}$ jako

$$f(x) = d(x, \phi(x)),\ \ x \in X$$

a ukažme, že f splňuje předpoklady Ekelandova variačního principu 2.1.3. Zřejmě je omezená zdola nulou. Protože je f složením dvou spojitých zobrazení, je spojitá.

Zvolme $\varepsilon \in (0, 1 - k)$, kde k je konstanta kontraktivity ϕ. Potom podle věty 2.4.1 existuje $y \in X$ takové, že pro všechna $x \in X$ platí nerovnost $f(y) - \varepsilon d(x, y) \leq f(x)$. Pokud dosadíme $x = \phi(y)$, dostáváme

$$d(y, \phi(y)) - \varepsilon d(\phi(y), y) \leq d(\phi(y), \phi^2(y)) \leq kd(y, \phi(y)),$$

tedy

$$(1 - \varepsilon - k)d(y, \phi(y)) \leq 0.$$

Platí tedy $d(y, \phi(y)) = 0$ a y je pevným bodem kontrakce.

Pokud by existovaly dva pevné body x a y, pak pro jejich vzdálenost platí

$$\begin{aligned} d(x, y) = d(\phi(x), \phi(y)) &\leq kd(x, y), \\ (1 - k)d(x, y) &\leq 0, \end{aligned}$$

a tedy $d(x, y) = 0$, což znamená $x = y$.

$\square$

Poznámka 2.4.5. Nechť jsou pro $(X, d), \Phi$ splněny předpoklady předchozí věty. Nechť kontrakce Φ je definována pomocí konstanty k. Vezměme libovolný bod $x_0 \in X$ a pro $i \in \mathbb{N}$ položme $x_i = \Phi(x_{i-1})$. Dokažme, že $\lim x_i$ existuje a tato limita je pevným bodem kontrakce.

Pro libovolné $m, n \in \mathbb{N},\ m > n$ platí

$$d(x_n, x_m) \leq \sum_{i=n}^{m-1} d(x_i, x_{i+1}) \leq \sum_{i=n}^{m-1} k^i d(x_0, x_1) \leq \sum_{i=n}^{\infty} k^i d(x_0, x_1) = \frac{k^n}{1 - k} d(x_0, x_1),$$

což znamená, že posloupnost $\{x_i\}$ je cauchyovská a díky úplnosti X konverguje k nějakému $x \in X$.

Protože Φ je spojitá, platí $\Phi(x_n) \to \Phi(x)$. Protože $\Phi(x_n) = x_{n+1}$, je x pevným bodem kontrakce.

Lemma 2.4.6. *Nechť $a < b$ jsou reálná čísla a nechť $f : (a,b) \to \mathbb{R}$ je spojitá funkce. Nechť dále $x, y, x', y' \in (a,b)$. Nechť X je graf funkce f na (a,b), tedy*

$$X = \{[x_1, x_2] \in \mathbb{R}^2; x_1 \in (a,b), x_2 = f(x_1)\}$$

a uvažujme na X metriku d. Nechť dále $\Phi : X \to X$ je spojité zobrazení, pro které platí $\Phi[x, f(x)] = [y, f(y)]$ a $\Phi[x', f(x')] = [y', f(y')]$.

Potom pro všechny $d \in \mathbb{R}$ v intervalu s krajními body y, y' existuje $c \in \mathbb{R}$ ležící mezi body x, x', takové, že $\Phi[c, f(c)] = [d, f(d)]$.

Důkaz. Definujme funkci $\Phi' : (a,b) \to \mathbb{R}$ takovou, že $\Phi'(h) = (\Phi[h, f(h)])_1$, kde $(.)_1$ udádá první souřadnici bodu z $\mathbb{R}^2$ a dokažme, že tato funkce je spojitá na (a,b). Pro $h_n \to h$ pak platí

$$\begin{aligned} \Phi[h_n, f(h_n)] &\to \Phi[h, f(h)], \\ \Phi([h_n, f(h_n)])_1 &\to \Phi([h, f(h)])_1, \\ \Phi'(h_n) &\to \Phi'(h). \end{aligned}$$

Tímto jsme problém převedli do $\mathbb{R}$. Protože $\Phi'(x) = y$ a $\Phi'(x') = y'$, pro všechny body d ležící v intervalu s krajními body y a y' existuje bod c ležící v intervalu s krajními body x a x' takový, že platí $\Phi'(c) = d$. To je však ekvivalentní s $\Phi[c, f(c)] = [d, f(d)]$, čímž je důkaz hotov.

□

Věta 2.4.7. *Existuje metrický prostor (X, d), který není úplný, ale každá kontrakce na něm má pevný bod.*

Důkaz. Definujme po částech lineární funkci $f : (0,1] \to \mathbb{R}$, která bude pro $k \in \mathbb{N}$ definována následovně:

$$f(\frac{1}{k}) = 0, f(\frac{k+\frac{1}{2}}{k(k+1)}) = (-1)^k.$$

Za prostor X pak zvolíme graf funkce f na $(0,1]$, tedy množinu

$$X = \{[x,y]; x \in (0,1], y = f(x)\}.$$

Metrika d na X bude euklidovská metrika v $\mathbb{R}^2$.

Na X zavedeme uspořádání $<$ následovně: $a < b \Leftrightarrow (a)_1 < (b)_1$, kde $(x)_1$ je první skouřadnice bodu $x \in X$. Podobně definujme uspořádání $>$, $\leq$ a $\geq$.

Mějme libovolné body $a, b \in X,\ a < b$. Řekneme, že $d_0, ..., d_{n+1} \in X$ je dělení mezi body a, b, pokud je splněno:

(i) $a = d_0,\ b = d_{n+1}$,

(ii) $d_0 < d_1 < ... < d_n < d_{n+1}$,

(iii) $|(d_i)_2| = 1$ pro všechna $i = 1, ..., n$,

(iv) neexistuje $y \in X \backslash \{d_1, ..., d_n\}$ splňující $a < y < b$ a $|(y)_2| = 1$.

Je zřejmé, že pro každé $a < b$ existuje právě jedno dělení mezi body a, b.

Pro $a < b$ najděme dělení $d_0, ..., d_{n+1}$ mezi body a, b. Potom definujme funkci $d' : X \times X \to [0, \infty)$ předpisem

$$d'(a, b) = \sum_{i=0}^{n} d(d_i, d_{i+1}),\ a < b.$$

Pokud $a > b$, položme $d'(a, b) = d'(b, a)$ a při $a = b$ položme $d'(a, b) = 0$. Potom d' je metrika na $X \times X$. Z trojúhelníkové nerovnosti pro d plyne pro všechny $a, b \in X$ nerovnost $d(a, b) \leq d'(a, b)$.

Mějme libovolnou kontrakci $\Phi : X \to X$ s konstantou kontraktivity k a ukažme, že pro libovolné body $a, b \in X$ platí nerovnost

$$d'(\Phi(a), \Phi(b)) \leq kd'(a, b).$$

Kvůli symetrii můžeme předpokládat $a < b$. Pokud $\Phi(a) = \Phi(b)$, je nerovnost splněna triviálně. Předpokládejme $\Phi(a) < \Phi(b)$. Pokud by platila obrácená nerovnost, důkaz se provede obdobně. Nechť $d_0, ..., d_{n+1}$ je dělení mezi body $\Phi(a), \Phi(b)$. Pro $i = 0, ..., n+1$ položme

$$M_i = \{x \in X; a \leq x \leq b, \Phi(x) = d_i\}$$

a

$$m_i = \inf\{(x)_1; x \in M_i\} \in [0, 1],$$
$$z_i = [m_i, f(m_i)] \in [a, b].$$

Z lemmatu 2.4.6 pak plyne, že M_i je neprázdná pro každé $i = 0, ..., n+1$.

Ukažme, že pro všechna $i = 0, ..., n$ platí $m_i \leq m_{i+1}$. Protože $m_0 = (a)_1$, platí $m_0 \leq m_1$. Nechť pro nějaké $i \in \{1, ..., n\}$ platí $m_i > m_{i+1}$. Použitím lemmatu 2.4.6

na d_0, d_i, d_{i+1} dostaneme, že existuje $x \in (a, m_{i+1})$ splňující $\Phi(x) = d_i$. Platí tedy $x < m_{i+1} < m_i$, což je spor s definicí m_i. Z toho hned plyne, že platí $z_i \leq z_{i+1}$.

Potom platí

$$\begin{aligned} d'(\Phi(a), \Phi(b)) &= \sum_{i=0}^{n} d(d_i, d_{i+1}) = \sum_{i=0}^{n} d(\Phi(z_i), \Phi(z_{i+1})) \leq k \sum_{i=0}^{n} d(z_i, z_{i+1}) \\ &\leq k \sum_{i=0}^{n} d'(z_i, z_{i+1}) = kd'(z_0, z_{n+1}) \leq kd'(a, b). \end{aligned}$$

Nechť $\hat{X}$ je zúplnění prostoru X v $\mathbb{R}^2$. K dokončení důkazu stačí ukázat, že pevný bod kontrakce Φ nemůže ležet v $\hat{X} \backslash X$, nemůže mít tedy první souřadnici nulovou. Zvolme $x_0 = [1, 0]$ a pro $n \in \mathbb{N}$ položme $x_n = \Phi(x_{n-1})$. Potom

$$\sum_{i=0}^{\infty} d'(x_i, x_{i+1}) \leq \sum_{i=0}^{\infty} k^i d'(x_0, x_1) = \frac{1}{1-k} d'(x_0, x_1) < \infty.$$

Podle poznámky 2.4.5 existuje $x = \lim x_n$ a x je pevným bodem kontrakce Φ. Pro libovolné $k \in \mathbb{N}$ platí

$$d'([1, 0], [\frac{1}{k}, 0]) = \sum_{i=1}^{k-1} d'([\frac{1}{i}, 0], [\frac{1}{i+1}, 0]) \geq \sum_{i=1}^{k-1} 2 = 2(k-1). \tag{2.11}$$

Pokud by platilo $x \notin X$, pak limitním přechodem $k \to \infty$ v (2.11) bychom dostali

$$\sum_{i=0}^{\infty} d'(x_i, x_{i+1}) \geq \lim_{i \to \infty} d'(x_0, x_i) = \infty,$$

což je spor. Tedy $x \in X$ a věta je dokázána.

□

Literatura

[1] Aliprantis C. D., Tourky R.: *Cones and Duality*, American Mathematical Society, Providence, Rhode Island, 2007.

[2] Borwein Jonathan M., Zhu Qiji J.: *Techniques of Variational Analysis*, Springer, New York, 2005.

[3] Franklin Joel: *Methods of Mathematical Economics*, Springer, New York, 1980.

[4] Lachout Petr: *Matematické programování*, preprint.

[5] Lindenstrauss Joram, Preiss David, Tišer Jaroslav: *A variational principle*, preprint.

[6] Rudin Walter: *Functional Analysis*, McGraw-Hill, 1973.