

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Jiří Krtek

Konvexní funkce a jejich zobecnění

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Milan Hladík, Katedra aplikované matematiky

Studijní program: Matematika, Obecná matematika

2006

Děkuji svému vedoucímu Mgr. Milanu Hladíkovi za veškeré rady a odborné vedení při psaní této bakalářské práce a rovněž za poskytnutí studijních materiálů.

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

V Praze dne 30.5.2006

Jiří Krtek

Obsah

1	Konvexní funkce	7
1.1	Základní vlastnosti konvexních funkcí	7
1.2	Diferencovatelnost konvexních funkcí	9
1.3	Minimalizace a maximalizace konvexních funkcí	12
2	Zobecněné konvexní funkce	13
2.1	Kvazikonvexní a explicitně kvazikonvexní funkce	13
2.2	Pseudokonvexní funkce	22
2.3	Minimalizace zobecněných konvexních a konkávních funkcí .	25
2.4	Některé další koncepty	27
	Literatura	36

Název práce: Konvexní funkce a jejich zobecnění
Autor: Jiří Krtek
Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: Mgr. Milan Hladík, Katedra aplikované matema-
tiky
e-mail vedoucího: milan_hladik@centrum.cz

Abstrakt: V předložené práci studujeme vlastnosti a vztahy mezi konvexními funkcemi a jejich zobecněními. Začínáme definicí konvexních funkcí a přes základní vlastnosti, jako je spojitost, se dostáváme k diferencovatelnosti a hledání jejich extrémů. Pokračujeme pojednáním o kvazikonvexních, explicitně kvazikonvexních a pseudokonvexních funkcích. Přes jejich definice a základní vlastnosti se dostáváme ke vztahům mezi nimi a konvexními funkcemi. Nalezneme zde i věty o skládání těchto zobecnění, které nám umožňují snadněji poznat, jestli je daná složená funkce (explicitně) kvazikonvexní, či pseudokonvexní. Práce obsahuje i část věnovanou minimalizaci těchto zobecnění. Na závěr práce jsou zmíněna některá další zobecnění konvexity, která už nejsou tak podrobně rozebrána.

Klíčová slova: konvexní, kvazikonvexní a pseudokonvexní funkce.

Title: Convex functions and their generalizations
Author: Jiří Krtek
Department: Department of Probability and Mathematical Statistics
Supervisor: Mgr. Milan Hladík, Department of applied mathematics
Supervisor's e-mail address: milan_hladik@centrum.cz

Abstract: In the present work we study properties and relations between convex functions and their generalizations. We commence with definition of convex functions and we get to differentiability and searching for extreme points through basic properties as continuity. We continue with quasiconvex, explicitly quasiconvex and pseudoconvex functions. Through their definitions and basic properties we get to relations between them and convex functions. We can find even theorems about composition of these generalizations here, which enable us easier to find, whether given composite function is (explicitly) quasiconvex or pseudoconvex. This work also contains a section dedicated to minimalization of these generalizations. There are mentioned some other generalizations of convexity at the conclusion of this work, which aren't analyzed so much.

Keywords: convex, quasiconvex and pseudoconvex functions.

Úvod

Záměrem této práce je představit případnému čtenáři konvexní a především zobecněné konvexní funkce. Práce se snaží shrnout veškeré vlastnosti jednotlivých typů těchto funkcí a také se snaží vykreslit vztahy mezi nimi. V nelineárním programování (dále zkracováno na NLP), což je jedna z disciplín optimalizace, se tyto typy funkcí často využívají.

Konvexní funkce mají spoustu silných vlastností, které ovšem občas nejsou zapotřebí. Zobecněné konvexní funkce tedy rozšiřují třídu konvexních funkcí tak, aby se zachovaly vlastnosti důležité pro optimalizaci. Zobecněným konvexním funkcím často chybí jen ty vlastnosti, které ani nepotřebujeme. To nám umožňuje řešit některé úlohy NLP za snížených předpokladů.

Úloha nalézt minimum nebo maximum nějaké funkce na nějaké množině se v optimalizaci objevuje velmi často. Řešitelnost těchto úloh popisují například Karushovy-Kahnovy-Tuckerovy a Fritz Johnovy podmínky optimality, ve kterých vystupují právě kvazikonvexní a pseudokonvexní funkce. Podmínky řešitelnosti takovýchto úloh tato práce neobsahuje. Po prostudování této práce však bude čtenář vědět, s jakými typy funkcí pracuje, když narazí na některou z výše zmíněných podmínek optimality.

V této práci jsou popisovány jen konvexní a zobecněné konvexní funkce, konkávní funkce a jejich zobecnění jsou zmiňovány jen v poznámkách a pak až v části věnované minimalizaci zobecněných konvexních a konkávních funkcí. Věty pro konkávní funkce a jejich zobecnění bychom dostali modifikací vět uvedených v tomto textu například změnou znaménka nebo obrácením nerovnosti, proto nejsou v tomto textu uvedeny. Navíc je to standardní postup většiny skript a vědeckých textů, vždy je popisován jen jeden z těchto konceptů.

Na začátku našeho textu stojí samozřejmě definice konvexních funkcí a pár jejich základních vlastností. Následují věty o diferencovatelnosti konvexních funkcí a s ním související téma hledání extrémů konvexních funkcí. To by bylo vše z první kapitoly. Druhá kapitola je mnohem obsáhlejší. Na začátku jsou uvedeny definice kvazikonvexních a explicitně kvazikonvexních funkcí, pár jejich základních vlastností, jejich vzájemné vztahy a jejich vztahy ke konvexním funkcím. Pokračujeme větami o složených kvazikonvexních, explicitně kvazikonvexních a konvexních funkcích. Rozebereme i lokální kvazikonvexitu a konvexitu. Tyto vlastnosti dovedou naše bádání rovnou k pseudokonvexním funkcím. Schéma části věnované pseudokonvexním funkcím je stejné jako schéma předchozí části, tj. nejdříve je uvedena definice a pár základních vlastností, následují vztahy k ostatním zobecněním a vztah ke konvexním funkcím. Poté shrneme veškeré vzájemné vztahy dosud zmíněných typů funkcí. Na závěr této části opět nechybí věty o složených pseudokonvexních funkcích. Následuje část věnovaná minimalizaci zobecněných konvexních funkcí. Poslední část druhé kapitoly, a vlastně poslední část této práce vůbec, obsahuje koncepty o kvazimonotónních, pseudomonotónních a ω -konvexních funkcích. Kvazimonotónní a pseudomonotónní funkce

zobecňují funkce lineární, ω -konvexní funkce spojují konvexní funkce s explicitně kvazikonvexními a kvazikonvexními.

Každé lokální minimum konvexní funkce je též globálním minimem. Toto je jedna z dobrých vlastností konvexních funkcí pro optimalizaci, která se s přechodem k explicitně kvazikonvexním a pseudokonvexním funkcím zachovává, kvazikonvexní funkce tuto vlastnost nemají. U pseudokonvexních funkcí je navíc tato vlastnost umocněna tím, že tyto funkce jsou z definice diferencovatelné, takže k nalezení globálního minima stačí nalézt stacionární bod. Naproti tomu derivace pseudokonvexních funkcí není neklesající, jako tomu je u konvexních funkcí. Explicitně kvazikonvexní funkce nejsou ani spojité na vnitřku konvexní množiny, jako tomu je u funkcí konvexních. Kvazikonvexní funkce jsou zobecněním explicitně kvazikonvexních funkcí. Množina bodů, ve kterých konvexní funkce nabývá svého minima na konvexní množině, je konvexní. To je jedna z dobrých vlastností konvexních funkcí, která se přechodem ke kvazikonvexním funkcím zachovává.

Při psaní této práce jsem nejvíce čerpal z [4] a z [1]. Explicitně kvazikonvexní funkce se definují různými způsoby, já jsem zvolil ten uvedený v [4], protože se mi zdál nejlogičtější. Dalším důvodem proč jsem zvolil tento koncept, je i ten fakt, že v knize [4] je uvedeno zdaleka nejvíc vět o zobecněných konvexních funkcích. To je zároveň i důvod, proč jsem z této knihy čerpal nejvíc.

To by bylo na úvod vše. Doufám, že se tato práce bude čtenáři líbit a že v ní najde potřebné informace.

Kapitola 1

Konvexní funkce

1.1 Základní vlastnosti konvexních funkcí

Tento text předpokládá znalosti základních matematických pojmů z následujících oblastí: metrických prostorů, diferenciálního počtu, konvexní analýzy, atd.

V následujícím textu budeme předpokládat $H \subset R^n$ konvexní a neprázdná s alespoň dvěma různými body, pokud nebude řečeno jinak. Dále $f : H \rightarrow R$ je funkce na konvexní množině H . Nejdříve tedy definujeme konvexní funkce.

Definice 1.1 *Nechť $f : H \rightarrow R$ je funkce definovaná na neprázdne konvexní množině H . Řekneme, že f je konvexní, jestliže pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ a $\lambda \in (0, 1)$ platí:*

$$f[(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2] \leq (1 - \lambda)f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2) \quad (1.1)$$

Dále definujeme:

- *Funkce je ryze konvexní, pokud ve vztahu (1.1) platí ostrá nerovnost.*
- *Funkce je konkávní, pokud ve vztahu (1.1) platí opačná nerovnost.*
- *Funkce je ryze konkávní, pokud ve vztahu (1.1) platí opačná ostrá nerovnost.*

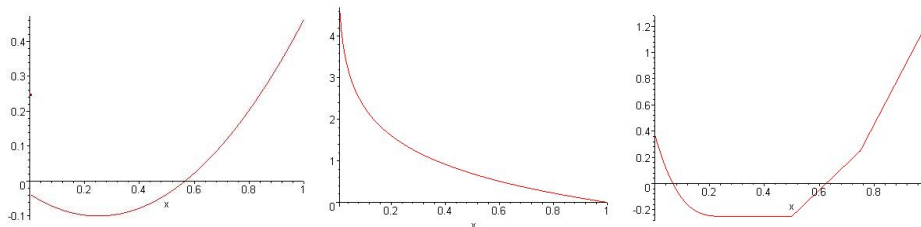
[1, strana 79].

K následující větě budeme potřebovat vědět, co je to epigraf dané funkce, a tak si ho zdefinujeme.

Definice 1.2 *Nechť $f : H \rightarrow R$ je funkce definovaná na neprázdne konvexní množině H . Epigraf funkce f , značení $\text{epi } f$, je podmnožina R^{n+1} definovaná jako:*

$$\text{epi } f = \{(\mathbf{x}, y) : \mathbf{x} \in H, y \in R, f(\mathbf{x}) \leq y\}.$$

[1, strana 84].



Obrázek 1.1: Grafy konvexních funkcí na intervalu $[0, 1]$

Následující věta se někdy používá, jako definice konvexních funkcí.

Věta 1.3 *Nechť $f : H \rightarrow R$ je funkce definovaná na konvexní množině H . Potom f je konvexní právě tehdy, když $\text{epi } f$ je konvexní množina.*

Důkaz viz [1, strana 84].

Definovali jsme si sice i funkce konkávní, ale budeme zde uvádět věty jen pro funkce konvexní. Věty pro konkávní funkce jsou naprosto analogické, tedy s pár změnami v nerovnostech, znaménkách atd. Podíváme se nyní na pár základních tvrzení o konvexních funkcích.

Lemma 1.4 *Nechť $f : H \rightarrow R$ je konvexní funkce na neprázdne konvexní množině H , pak je množina*

$$L_\alpha = \{\mathbf{x} \in H : f(\mathbf{x}) \leq \alpha\}$$

konvexní.

Důkaz viz [1, strana 80].

Věta 1.5 *Nechť $f : H \rightarrow R$ je konvexní funkce na neprázdne konvexní množině H , potom f je spojitá na $\text{int}(H)$, tj. na vnitřku množiny H .*

Důkaz viz [1, strana 81].

Věta 1.6 *Nechť H je neprázdne konvexní množina a $f_1 : H \rightarrow R, \dots, f_k : H \rightarrow R$ jsou konvexní funkce na H . Pak pro všechna $\alpha_1 \geq 0, \dots, \alpha_k \geq 0$ je $\sum_{i=1}^k \alpha_i f_i$ opět konvexní funkce.*

Důkaz viz [4, strana 60].

Věta 1.7 (Jensenova nerovnost I.) *Nechť $f : H \rightarrow R$ je konvexní funkce na neprázdne konvexní množině H , pak pro každé $n \in N$, pro všechna $\mathbf{x}_1, \dots, \mathbf{x}_n \in H$ a všechna $\alpha_1 \geq 0, \dots, \alpha_n \geq 0$ taková, že $\alpha_1 + \dots + \alpha_n = 1$, platí:*

$$f\left(\sum_{k=1}^n \alpha_k \mathbf{x}_k\right) \leq \sum_{k=1}^n f(\alpha_k \mathbf{x}_k). \quad (1.2)$$

Důkaz viz [5, strana 89].

Věta 1.8 (Jensenova nerovnost II.) *Nechť $\mathbf{X} = (X_1, \dots, X_k)^t$ je reálný náhodný vektor s hodnotami v neprázdné konvexní množině $D \subset R^k$ a necht' mají náhodné veličiny X_1, \dots, X_k konečnou střední hodnotu. Pak $E[\mathbf{X}] = (E[X_1], \dots, E[X_k])^t \in D$ a pro každou konvexní funkci $G : D \rightarrow R$ je buď $E[G(\mathbf{X})^+] < +\infty$, nebo $E[G(\mathbf{X})^-] < +\infty$ a navíc $G(E[\mathbf{X}]) \leq E[G(\mathbf{X})]$.*

Důkaz viz [3, strana 26].

Věta 1.9 *Nechť $f : R^n \rightarrow R$ je funkce, definujme pro každý bod $\mathbf{x} \in R^n$ a pro každý nenulový směr $\mathbf{d} \in R^n$ $F_{(\mathbf{x};\mathbf{d})}(\lambda) = f(\mathbf{x} + \lambda\mathbf{d})$ jako funkci proměnné $\lambda \in R$. Potom je f (ryze) konvexní právě tehdy, když F je (ryze) konvexní pro všechna \mathbf{x} a $\mathbf{d} \neq 0$, $\mathbf{x}, \mathbf{d} \in R^n$.*

Důkaz viz [1, strana 94].

1.2 Diferencovatelnost konvexních funkcí

Už víme, že na vnitřku konvexní množiny (proto také na otevřené konvexní množině) je konvexní funkce spojitá. Nyní si ukážeme, jak je to s diferencovatelností konvexní funkce na konvexní množině. Začneme derivací ve směru. K tomu budeme potřebovat následující definici.

Definice 1.10 *Nechť $H \subset R^n$ je neprázdná a $f : H \rightarrow R$ je funkce definovaná na množině H . Necht' $\mathbf{x}_0 \in H$ a \mathbf{d} je nenulový vektor takový, že $\mathbf{x}_0 + \lambda\mathbf{d} \in H$ pro nějaké $\lambda > 0$ dostatečně malé. Derivace funkce f v bodě \mathbf{x}_0 ve směru \mathbf{d} , značení $f'(\mathbf{x}_0; \mathbf{d})$, je definována následující limitou*

$$f'(\mathbf{x}_0; \mathbf{d}) = \lim_{\lambda \rightarrow 0^+} \frac{f(\mathbf{x}_0 + \lambda\mathbf{d}) - f(\mathbf{x}_0)}{\lambda},$$

pokud existuje.

[1, strana 82].

Lemma 1.11 *Nechť $f : R^n \rightarrow R$ je konvexní funkce. Necht' $\mathbf{x}_0 \in R^n$ a $\mathbf{d} \in R^n$ je nenulový vektor. Potom derivace $f'(\mathbf{x}_0; \mathbf{d})$ funkce f v bodě \mathbf{x}_0 ve směru \mathbf{d} existuje.*

Důkaz viz [1, strana 83].

Abychom mohli v tomto tématu pokračovat, musíme zavést derivaci v n -rozměrném reálném prostoru, tzv. totální diferenciál, a s ním související pojem gradientu.

Definice 1.12 *Nechť $f : R^n \rightarrow R$ je funkce n proměnných, $\mathbf{a} \in R^n$. Necht' $L : R^n \rightarrow R$ je lineární forma. Řekneme, že L je totální diferenciál funkce f v bodě \mathbf{a} (nebo také derivace funkce f v bodě \mathbf{a}), jestliže platí:*

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

Totální diferenciál L se někdy značí $f'(\mathbf{a})$ a jakožto lineární forma má tedy tvar:

$$L(h_1, \dots, h_n) = A_1 h_1 + \dots + A_n h_n,$$

kde $A_1, \dots, A_n \in \mathbb{R}$.

[6, strana 52].

Poznámka 1.13 Má-li funkce totální diferenciál, pak říkáme, že je diferencovatelná. Poznamenejme ještě, že pokud totální diferenciál funkce f v bodě \mathbf{a} existuje, pak je určen jednoznačně a má tvar:

$$L(\mathbf{h}) = \frac{\partial f(\mathbf{a})}{\partial x_1} h_1 + \dots + \frac{\partial f(\mathbf{a})}{\partial x_n} h_n$$

kde $\frac{\partial f(\mathbf{a})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{a})}{\partial x_n}$ jsou parciální derivace funkce f v bodě \mathbf{a} a $\mathbf{h} = (h_1, \dots, h_n)^t \in \mathbb{R}^n$.

S diferencovatelností funkce souvisí již zmíněný gradient. Pokud existuje gradient dané funkce v daném bodě, pak je tato funkce v tomto bodě diferencovatelná. Gradient je vlastně vektor (směr) největšího stoupání dané funkce v daném bodě. Jestliže tedy existuje gradient funkce f v bodě \mathbf{a} pak má tvar:

$$\nabla f(\mathbf{a}) = \left(\frac{\partial f(\mathbf{a})}{\partial x_1}, \dots, \frac{\partial f(\mathbf{a})}{\partial x_n} \right)^t.$$

Nyní si uvedeme několik vět o diferencovatelnosti konvexních funkcí.

Věta 1.14 Necht H je neprázdná otevřená konvexní množina v \mathbb{R}^n a $f : H \rightarrow \mathbb{R}$ je funkce diferencovatelná na H . Potom f je konvexní tehdy a jen tehdy, když pro každý bod $\mathbf{x}_0 \in H$ a pro všechna $\mathbf{x} \in H$ platí:

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^t (\mathbf{x} - \mathbf{x}_0).$$

Podobně, funkce f je ryze konvexní tehdy a jen tehdy, když pro každý bod $\mathbf{x}_0 \in H$ a pro všechna $\mathbf{x} \in H, \mathbf{x} \neq \mathbf{x}_0$ platí:

$$f(\mathbf{x}) > f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^t (\mathbf{x} - \mathbf{x}_0).$$

Důkaz viz [5, strana 98].

Věta 1.15 Necht H je neprázdná otevřená konvexní množina v \mathbb{R}^n a $f : H \rightarrow \mathbb{R}$ je funkce diferencovatelná na H . Potom f je konvexní tehdy a jen tehdy, když pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ platí:

$$[\nabla f(\mathbf{x}_2) - \nabla f(\mathbf{x}_1)]^t (\mathbf{x}_2 - \mathbf{x}_1) \geq 0.$$

Podobně, f ryze konvexní tehdy a jen tehdy, když pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ různá platí:

$$[\nabla f(\mathbf{x}_2) - \nabla f(\mathbf{x}_1)]^t (\mathbf{x}_2 - \mathbf{x}_1) > 0.$$

Důkaz viz [1, strana 90].

Věta 1.16 *Nechť H je otevřená množina v R^n a $f : H \rightarrow R$ je konvexní funkce. Má-li f parciální derivace v bodě $\mathbf{x} \in H$, pak má v bodě \mathbf{x} i totální diferenciál.*

Důkaz viz [5, strana 101]

Dále budeme potřebovat vědět, jak vypadá druhá derivace funkce n-proměnných. To ukazuje další definice.

Definice 1.17 *Nechť H je neprázdná množina v R^n a $f : H \rightarrow R$ je funkce na H . Potom je funkce f dvakrát diferencovatelná v bodě $\mathbf{x}_0 \in \text{int}(H)$, jestliže existuje vektor $\nabla f(\mathbf{x}_0)$, $n \times n$ symetrická matice $\mathbf{H}(\mathbf{x}_0)$, matice druhých derivací funkce f , a funkce $\alpha : R^n \rightarrow R$ tak, že platí*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^t (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^t \mathbf{H}(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) + \|\mathbf{x} - \mathbf{x}_0\| \alpha(\mathbf{x}_0; \mathbf{x} - \mathbf{x}_0)$$

pro každé $\mathbf{x} \in H$, kde $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \alpha(\mathbf{x}_0; \mathbf{x} - \mathbf{x}_0) = 0$. Řekneme, že je funkce f dvakrát diferencovatelná na otevřené množině $H_0 \subset H$, jestliže je dvakrát diferencovatelná v každém bodě množiny H_0 . Matice druhých derivací funkce f je složena z druhých parciálních derivací funkce f $\partial^2 f / \partial x_i \partial x_j$ $i, j = 1, \dots, n$ a má tvar:

$$\mathbf{H}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_n^2} \end{pmatrix}$$

[1, strana 90].

Teď už jsme tedy připraveni nahlédnout další vlastnosti konvexních funkcí. Ukážeme si, jak snadněji poznat, že je funkce konvexní. Budeme předpokládat, že daná funkce je dvakrát diferencovatelná na dané množině, a pomocí matice druhých derivací pak určíme, zda-li je funkce konvexní nebo ne.

Věta 1.18 *Nechť H je neprázdná otevřená konvexní množina v R^n a funkce $f : H \rightarrow R$ je dvakrát diferencovatelná na H . Potom je funkce f konvexní právě tehdy, když matice druhých derivací je pozitivně semidefinitní v každém bodě množiny H . Pokud je navíc matice druhých derivací funkce f pozitivně definitní v každém bodě množiny H , pak je funkce f ryze konvexní.*

Důkaz viz [5, strana 100].

Věta 1.19 *Nechť H je neprázdná otevřená konvexní množina v R a $f : H \rightarrow R$ má derivace všech řádů. Potom f je ryze konvexní na H tehdy a jen tehdy když, pro každý bod $x \in H$ existuje $n \in N$ takové, že $f^{(n)}(x) > 0$ a $f^{(j)}(x) = 0$ pro všechna $1 < j < n$, kde $f^{(n)}$ značí j -tou derivaci funkce f .*

Důkaz viz [1, strana 93].

1.3 Minimalizace a maximalizace konvexních funkcí

V této části si ukážeme využití konvexních funkcí v optimalizaci. Uvedeme si pár vět týkajících se nabývání minima, či maxima konvexní funkce na konvexní množině. Nebudeme definovat (ostré) globální, respektive globální, minimum, respektive maximum. Čtenář bude jistě tyto základní pojmy znát.

Věta 1.20 *Nechť H je neprázdná konvexní množina v R^n a $f : H \rightarrow R$ je konvexní funkce. Nechť $\mathbf{x}_0 \in H$ je bodem lokálního minima funkce f na množině H . Pak je \mathbf{x}_0 i bodem globálního minima funkce f na množině H . Pokud je navíc funkce f ryze konvexní nebo je \mathbf{x}_0 bodem ostrého lokálního minima, pak je toto globální minimum určeno jednoznačně.*

Důkaz viz [1, strana 101].

Věta 1.21 *Nechť H je otevřená konvexní množina v R^n a $f : H \rightarrow R$ je konvexní funkce. Jestliže $\nabla f(\mathbf{x}_0) = \mathbf{0}$ v bodě $\mathbf{x}_0 \in H$, potom $f(\mathbf{x}_0)$ je globální minimum funkce f . Je-li navíc f spojitě diferencovatelná v okolí bodu \mathbf{x}_0 a matice druhých derivací funkce f v bodě \mathbf{x}_0 je pozitivně definitní, pak je \mathbf{x}_0 jednoznačně určený bod globálního minima funkce f na množině H .*

Důkaz viz [5, strana 124].

Věta 1.22 *Buď $f : H \rightarrow R$ spojitá a konvexní funkce a $H \subset R^n$ kompaktní konvexní polyedrická množina. Potom funkce f nabývá svého maxima na hranici množiny H .*

Důkaz viz [1, strana 107].

Kapitola 2

Zobecněné konvexní funkce

V této kapitole si představíme další třídy funkcí, které nemají tolik dobrých vlastností jako funkce (ryze) konvexní, ale v mnoha důležitých vlastnostech se shodují. V optimalizaci často stačí předpokládat právě různé typy zobecněné konvexity, např.: kvazikonvexitu, explicitní kvazikonvexitu, či pseudokonvexitu, namísto konvexity. Začneme kvazikonvexními a explicitně kvazikonvexními funkcemi, které nevyžadují diferencovatelnost. Pokračovat budeme pseudokonvexními funkcemi a na závěr si uvedeme ještě nějaká další zajímavá zobecnění konvexity.

2.1 Kvazikonvexní a explicitně kvazikonvexní funkce

V celé této části budeme předpokládat, že $H \subset R^n$ je konvexní neprázdná množina s alespoň dvěma různými body, pokud nebude řečeno jinak.

Definice 2.1 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$. Řekneme, že funkce f je kvazikonvexní, jestliže pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ a všechna $\mathbf{x}_0 \in (\mathbf{x}_1, \mathbf{x}_2)$ platí:*

$$f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \Rightarrow f(\mathbf{x}_0) \leq f(\mathbf{x}_2). \quad (2.1)$$

Vztah 2.1 lze ekvivalentně vyjádřit ve tvaru:

$$f(\mathbf{x}_0) \leq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}. \quad (2.2)$$

[4, strana 44].

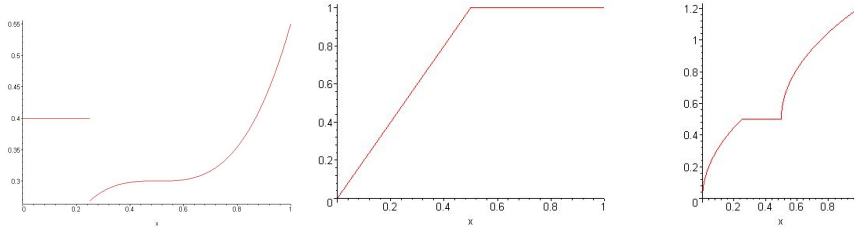
Definice 2.2 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$. Řekneme, že funkce f je explicitně kvazikonvexní, jestliže je kvazikonvexní a pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ a všechna $\mathbf{x}_0 \in (\mathbf{x}_1, \mathbf{x}_2)$ platí:*

$$f(\mathbf{x}_1) < f(\mathbf{x}_2) \Rightarrow f(\mathbf{x}_0) < f(\mathbf{x}_2). \quad (2.3)$$

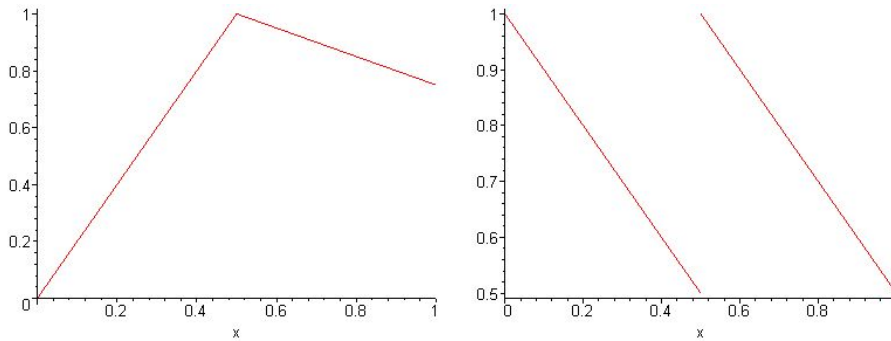
Vztah 2.3 lze také vyjádřit ve tvaru:

$$f(\mathbf{x}_1) \neq f(\mathbf{x}_2) \Rightarrow f(\mathbf{x}_0) < \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}, \quad (2.4)$$

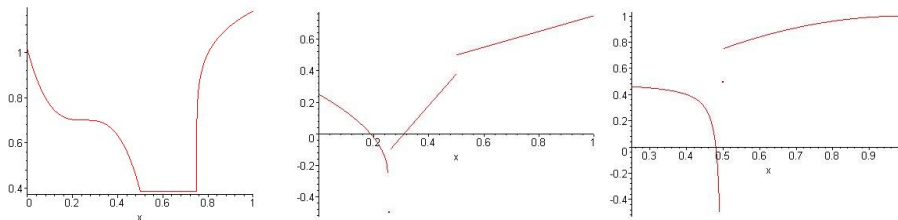
[4, strana 44].



Obrázek 2.1: Grafy kvazikonvexních funkcí na intervalu $[0, 1]$.



Obrázek 2.2: Grafy nekvazikonvexních funkcí na intervalu $[0, 1]$.



Obrázek 2.3: Grafy explicitně kvazikonvexních funkcí na intervalu $[0, 1]$

Poznámka 2.3 Funkce f je (explicitně) kvazikonkávní na konvertní neprázdné množině H , jestliže je funkce $-f$ (explicitně) kvazikonvexní na H .

Nejdříve si uvedeme pár jednoduchých tvrzení týkajících se vztahů mezi kvazikonvexními, explicitně kvazikonvexními a konvexními funkcemi.

Lemma 2.4 *Jestliže je funkce f konvexní na konvexní neprázdné množině H , potom je f explicitně kvazikonvexní na H . Obrácená implikace neplatí.*

Důkaz viz [4, strana 46].

Příklad 2.5 *Mějme funkci $f(x) = x^3$, $x \in \mathbb{R}$. Tato funkce je explicitně kvazikonvexní, protože pro všechna $x_1, x_2 \in \mathbb{R}$ a $x_0 \in (x_1, x_2)$ splňuje vztah (2.4). Ale není konvexní, protože $f'' < 0$ na intervalu $(-\infty, 0)$.*

Lemma 2.6 *Jestliže je funkce f explicitně kvazikonvexní na konvexní neprázdné množině H , potom je f kvazikonvexní na H . Obrácená implikace neplatí.*

Důkaz viz [4, strana 47].

Příklad 2.7 *Mějme funkci*

$$f(x) = \begin{cases} 0 & x \geq 0 \\ x & x < 0 \end{cases}$$

Tato funkce je jistě kvazikonvexní, neboť splňuje vztah (2.2). Ale není explicitně kvazikonvexní, stačí ve vztahu (2.4) položit $x_1 = -1$, $x_2 = 2$. Tím nám vztah (2.4) selže pro všechna $0 \leq x_0 < 2$.

Někdy se explicitně kvazikonvexní funkce definují bez požadavku kvazikonvexnosti. To ale vede k tomu, že některé funkce které jsou explicitně kvazikonvexní, nejsou kvazikonvexní. Tuto skutečnost ukazují následující příklad, známý pod anglickým názvem Karamardian's anomaly.

Příklad 2.8 *Definujme funkci $f : \mathbb{R}^n \rightarrow \mathbb{R}$ následovně:*

$$f(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \neq \mathbf{0} \\ 1 & \mathbf{x} = \mathbf{0} \end{cases}$$

Snadno zjistíme, že vlastnost (2.3) platí všude v \mathbb{R}^n . Avšak funkce f není kvazikonvexní, protože pro každý bod $\bar{\mathbf{x}} \neq \mathbf{0}$ je $\mathbf{0} \in (\bar{\mathbf{x}}, -\bar{\mathbf{x}})$ a platí:

$$1 = f(\mathbf{0}) > f(\bar{\mathbf{x}}) = f(-\bar{\mathbf{x}}) = 0,$$

což je ve sporu s (2.2), takže f nemůže být kvazikonvexní.

[4, strana 45].

Na tomto místě si zadefinujeme méně známý pojem dolní polospojivosti a hned za touto definicí si ukážeme, jak se vyhnout požadavku na kvazikonvexnost funkce v definici explicitní kvazikonvexnosti.

Definice 2.9 *Nechť H je neprázdná množina v \mathbb{R}^n a $f : H \rightarrow \mathbb{R}$ je funkce na H . Řekneme, že funkce f je zdola polospojita v bodě $\bar{\mathbf{x}} \in H$, jestliže pro každé $\varepsilon > 0$ existuje $\delta > 0$ tak, že všechna $\mathbf{x} \in H$, $\|\mathbf{x} - \bar{\mathbf{x}}\| < \delta$ platí $f(\mathbf{x}) - f(\bar{\mathbf{x}}) > -\varepsilon$.*

[1, strana 563].

Věta 2.10 *Nechť $f : H \rightarrow R$ je zdola polospojité funkce na neprázdné konvexní množině $H \subset R^n$ a nechť splňuje vztah (2.3) pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ a všechna $\mathbf{x}_0 \in (\mathbf{x}_1, \mathbf{x}_2)$. Potom je funkce f kvazikonvexní a tedy i explicitně kvazikonvexní na H .*

Důkaz viz [4, strana 45].

Nyní uvedeme větu, která se podobá lemmatu 1.4 a také někdy slouží jako definice kvazikonvexních funkcí.

Věta 2.11 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$. Potom je funkce f kvazikonvexní tehdy a jen tehdy, když je množina $S_\alpha = \{\mathbf{x} \in H; f(\mathbf{x}) \leq \alpha\}$ konvexní pro každé $\alpha \in R$.*

Důkaz viz [1, strana 108].

Další dvě věty říkají, jak vlastně kvazikonvexní a explicitně kvazikonvexní funkce vypadají, co se týče monotonie. Ještě než si je uvedeme, zavedeme si pro zjednodušení zápisu následující značení: $\langle \alpha, \beta \rangle$ bude znamenat, že nevíme zda-li do tohoto intervalu patří koncové body. Předpokládáme, že čtenář bude znát definice (ne)rostoucí a (ne)klesající funkce, a tak je tu neuvádíme.

Věta 2.12 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$. Potom je f kvazikonvexní právě tehdy, když pro každý segment $[\mathbf{x}_1, \mathbf{x}_2] \subset H$ platí jedna z následujících podmínek:*

- (i) *funkce f je nerostoucí nebo neklesající na $[\mathbf{x}_1, \mathbf{x}_2]$*
- (ii) *funkce f je nerostoucí na $(\mathbf{x}_1, \mathbf{x}_2]$, ale ne na $[\mathbf{x}_1, \mathbf{x}_2]$ nebo neklesající na $[\mathbf{x}_1, \mathbf{x}_2)$, ale ne na $[\mathbf{x}_1, \mathbf{x}_2]$*
- (iii) *existuje bod $\bar{\mathbf{x}} \in (\mathbf{x}_1, \mathbf{x}_2)$ tak, že je funkce f nerostoucí na $[\mathbf{x}_1, \bar{\mathbf{x}})$ a neklesající na $(\bar{\mathbf{x}}, \mathbf{x}_2]$, kde alespoň jeden z těchto segmentů je uzavřený.*

Důkaz viz [4, strana 49].

Věta 2.13 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$. Potom je f explicitně kvazikonvexní právě tehdy, když lze každý segment $[\mathbf{x}_1, \mathbf{x}_2] \subset H$ rozdělit na tři segmenty (každý dělicí bod musí být obsažen alespoň v jednom ze segmentů, které uzavírá) tak, že je funkce f klesající na prvním, konstantní na druhém a rostoucí na třetím segmentu. Jakýkoli jeden nebo i dva z těchto segmentů mohou být prázdné nebo degenerované do jednoho bodu. Na druhém segmentu, pokud není prázdný, musí funkce f nabývat svého minima.*

Důkaz viz [4, strana 51].

Lemma 2.14 *Nechť $f : H \rightarrow R$ je kvazikonvexní funkce na neprázdné konvexní množině $H \subset R$. Potom je množina bodů nespojitosti funkce f nejvýše spočetná.*

[4, strana 53].

Povšimněme si, že právě uvedené lemma neplatí pro $H \subset R^n$. Pro kvazikonvexní funkce více proměnných platí následující věta, která je vlastně zobecněním uvedeného lemmatu.

Věta 2.15 *Nechť $f : H \rightarrow R$ je kvazikonvexní funkce na neprázdne konvexní množině $H \subset R^n$. Potom je funkce f spojitá skoro všude v H .*

[4, strana 55].

Nyní se dostáváme k části věnované skládání kvazikonvexních, explicitně kvazikonvexních a konvexních funkcí. Občas bývá těžké určit, zda-li má daná funkce jednu z vlastností jmenovaných výše. Tuto funkci lze ovšem rozložit na jednodušší části, ze kterých je složena a podle nich rozhodnout, jestli je kvazikonvexní/konkávní, explicitně kvazikonvexní/konkávní, atd.

V následujícím textu budeme potřebovat vědět, co je konvexní zobrazení. Konvexním zobrazením se tedy rozumí, že každá z jeho složek je konvexní. Před vyslovením další věty pro zjednodušení zápisu nejdříve zdefinujeme, kdy je funkce (ryze) rostoucí/klesající v i -té složce.

Definice 2.16 *Nechť $f : H \rightarrow R$ je funkce na neprázdne konvexní množině $H \subset R^n$. Řekneme, že je funkce f rostoucí/klesající v i -té souřadnici pro nějaké $i \in \{1, \dots, n\}$, jestliže je pro každá $\mathbf{x}^1, \mathbf{x}^2 \in H$ taková, že $\mathbf{x}^2 - \mathbf{x}^1 = (\mathbf{x}_i^2 - \mathbf{x}_i^1)e_i \geq \mathbf{0}$, funkce f rostoucí/klesající v $[\mathbf{x}^1, \mathbf{x}^2]$. V této definici neznamená horní index mocninu, ale rozlišení jednotlivých bodů. Dolní index vyjadřuje souřadnici a e_i je i -tý vektor kanonické báze v R^n . Ryze rostoucí/klesající funkce v i -té souřadnici se definuje obdobně.*

[4, strana 57].

Věta 2.17 *Nechť $g : H \rightarrow R^m$ je konvexní zobrazení na konvexní neprázdne množině $H \subset R^n$ a nechť $f : \text{conv}[g(H)] \rightarrow R$ je funkce na konvexním obalu množiny $g(H) \subset R^m$. Pokud je funkce f rostoucí v každé souřadnici i , ve které není i -tá složka zobrazení g lineární, a je na $\text{conv}[g(H)]$*

- (i) kvazikonvexní
- (ii) explicitně kvazikonvexní
- (iii) konvexní,

potom je funkce $\varphi(\mathbf{x}) \equiv f[g(\mathbf{x})]$, $\mathbf{x} \in H$,

- (i) kvazikonvexní
- (ii) explicitně kvazikonvexní
- (iii) konvexní.

Důkaz viz [4, strana 58].

Různými změnami předpokladů na vlastnosti funkce f a zobrazení g v předchozí větě dosáhneme podobných výsledků. Tyto modifikace shrnuje tabulka 2.1.

V následující tabulce jsou použity tyto zkratky:

- rostoucí, respektive klesající, v NL znamená, že funkce f je rostoucí, respektive klesající, v každé souřadnici, ve které není i -tá složka zobrazení g lineární
- konv. znamená samozřejmě konvexní, podobně konk. znamená konkávní
- (ex.) znamená explicitně
- (kv.) znamená kvazi

Varianta	1	2	3	4
g	konvexní		konkávní	
f	neklesající v NL	nerostoucí v NL	neklesající v NL	nerostoucí v NL
f	(ex.)(kv.) konv.	(ex.)(kv.) konk.	(ex.)(kv.) konk.	(ex.)(kv.) konv.
$f \circ g$	(ex.)(kv.) konv.	(ex.)(kv.) konk.	(ex.)(kv.) konk.	(ex.)(kv.) konv.

Tabulka 2.1: Varianty věty 2.17

Věta 2.18 *Nechť $g : H \rightarrow R$ je kvazikonvexní funkce na konvexní neprázdné množině $H \subset R^n$ a nechť $f : g(H) \rightarrow R$ je neklesající funkce na množině $g(H) \subset R$. Potom je funkce $\varphi(\mathbf{x}) \equiv f[g(\mathbf{x})]$ kvazikonvexní na H . Je-li navíc g explicitně kvazikonvexní a f rostoucí, pak je φ explicitně kvazikonvexní.*

Důkaz viz [4, strana 59].

Z předchozí věty je vidět, že některé kvazikonvexní funkce dostaneme složením funkce konvexní (ve větě funkce g) s funkcí rostoucí (ve větě funkce f). Pokud k této rostoucí funkci existuje na nějakém intervalu inverzní funkce, pak je na tomto intervalu rovněž rostoucí. A proto některé kvazikonvexní funkce můžeme převést na konvexní složením s vhodnou rostoucí funkcí. Odpověď na otázku, jestli lze každou kvazikonvexní funkci takto převést na konvexní, je záporná. Existují kvazikonvexní i explicitně kvazikonvexní funkce, které nelze převést na konvexní.

Předchozí věta má rovněž různé varianty, které shrnují další tabulky 2.2 a 2.3.

Nyní si uvedeme pár jednoduchých aplikací a důsledků dvou předešlých vět, aby si čtenář udělal představu, k čemu je lze využít.

Lemma 2.19 *Funkce*

$$f(\mathbf{x}) \equiv (x_1)^{\alpha_1} (x_2)^{\alpha_2} \dots (x_n)^{\alpha_n},$$

kde $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$ a $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$, je explicitně kvazikonkávní na množině $R_+^n = \{\mathbf{x} \in R^n : x_i \geq 0, i = 1, \dots, n\}$.

Důkaz viz [4, strana 61].

Varianta	1	2	3	4
g	kvazikonvexní		kvazikonkávní	
f	neklesající	nerostoucí	neklesající	nerostoucí
$f \circ g$	kvazikonvexní	kvazikonkávní		kvazikonvexní

Tabulka 2.2: Varianty věty 2.18

Varianta	1	2	3	4
g	explicitně kvazikonvexní		explicitně kvazikonkávní	
f	rostoucí	klesající	rostoucí	klesající
$f \circ g$	expl. kvazikonvexní	expl. kvazikonkávní	expl. kvazikonvexní	

Tabulka 2.3: Varianty věty 2.18

Lemma 2.20 *Nechť $g : H \rightarrow R^m$ je konkávní a nezáporné zobrazení na konvexní množině $H \subset R^n$ a $\alpha_1, \alpha_2, \dots, \alpha_m \geq 0$. Potom je funkce*

$$f(\mathbf{x}) \equiv \prod_{i=1}^m [g_i(\mathbf{x})]^{\alpha_i}$$

explicitně kvazikonkávní na H .

Důkaz viz [4, strana 61].

Další tvrzení je přímým důsledkem předchozího lemmatu.

Důsledek 2.21 *Nechť $f, g : H \rightarrow R$ jsou dvě konkávní a nezáporné funkce na konvexní množině $H \subset R^n$. Potom je funkce $\varphi \equiv f \cdot g$ explicitně kvazikonkávní na H .*

[4, strana 61].

Varianta	1	2	3
g	konvexní, nekladná	konvexní, nekladná	konkávní, nezáporná
f	konvexní, nekladná	konkávní, nezáporná	konkávní, nezáporná
$f \cdot g$	expl. kvazikonkávní	expl. kvazikonvexní	expl. kvazikonkávní

Tabulka 2.4: Varianty důsledku 2.21

Lemma 2.22 *Nechť $f : H \rightarrow R$ je konkávní a kladná funkce na konvexní množině $H \subset R^n$. Potom je funkce $\varphi \equiv 1/f$ konvexní na množině H .*

Důkaz viz [4, strana 62].

Lemma 2.23 *Nechť $f, g : H \rightarrow R$ jsou funkce na konvexní množině $H \subset R^n$. Nechť je funkce f konvexní a nezáporná a funkce g ať je konkávní a kladná. Potom je funkce*

$$\varphi \equiv \frac{f}{g}$$

explicitně kvazikonvexní na množině H .

Důkaz viz [4, strana 62].

Varianta	1	2	3	4	5	6	7	8
f	konvexní		konkávní		konvexní		konkávní	
f	nezáporná	nekladná		nezáporná		nekladná	nezáporná	
g	konkávní	konvexní	konkávní	konvexní	konkávní	konvexní	konkávní	konvexní
g	kladná		záporná		záporná		kladná	
f/g	explicitně kvazikonvexní				explicitně kvazikonkávní			

Tabulka 2.5: Varianty lemmatu 2.23

Lemma 2.24 *Nechť $f : H \rightarrow R$ je konvexní funkce na konvexní množině $H \subset R^n$. Nechť je funkce $g(\mathbf{x}) \equiv \mathbf{a}^t \mathbf{x} + \alpha$ kladná na množině H . Potom je funkce*

$$\varphi(\mathbf{x}) \equiv \frac{f(\mathbf{x})}{\mathbf{a}^t \mathbf{x} + \alpha}$$

explicitně kvazikonvexní na množině H .

Důkaz viz [4, strana 63].

I toto lemma má různé varianty, které zde pro úplnost uvedeme. Shrnuje je tabulka 2.6.

Varianta	1	2	3	4
f	konvexní		konkávní	
$\mathbf{a}^t \mathbf{x} + \alpha$	kladná	záporná	kladná	záporná
$\frac{f(\mathbf{x})}{\mathbf{a}^t \mathbf{x} + \alpha}$	expl. kvazikonvexní	expl. kvazikonkávní	expl. kvazikonvexní	expl. kvazikonkávní

Tabulka 2.6: Varianty lemmatu 2.24

Věta 2.25 *Nechť $g : H \rightarrow R^m$ je*

- (i) *kvazikonvexní*
- (ii) *explicitně kvazikonvexní*

(iii) konvevní

zobrazení na konvevní neprázdné množině $H \subset R^n$. Potom je funkce $f \equiv \max\{g_1(\mathbf{x}), \dots, g_m(\mathbf{x})\}$

(i) kvazikonvevní

(ii) explicitně kvazikonvevní

(iii) konvevní

na množině H .

Důkaz viz [4, strana 64].

Nyní budeme definovat lokální kvazikonvexitu, kvazikonkávitu, konvexitu a konkávitu, což vyžaduje diferencovatelnost funkce v daném bodě, za to ale nepotřebujeme, aby byla množina, na které pojmy definujeme, konvevní. Uvedeme pár vět týkajících se vztahu lokální (kvazi)konvexity a (kvazi)konvexity na množině.

Definice 2.26 Necht' $f : H \rightarrow R$ je funkce na neprázdné množině $H \subset R^n$ a necht' je diferencovatelná v bodě $\bar{\mathbf{x}} \in H$. Řekneme, že je funkce f lokálně

(i) kvazikonvevní

(ii) konvevní

v bodě $\bar{\mathbf{x}}$, jestliže pro všechna $\mathbf{x} \in H$ platí:

$$(i) \quad f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) \Rightarrow (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) \leq 0 \quad (2.5)$$

$$(ii) \quad f(\mathbf{x}) - f(\bar{\mathbf{x}}) - (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) \geq 0 \quad (2.6)$$

Lokální kvazikonvexita viz [4, strana 108], lokální konvexita viz [4, strana 110].

Můžeme poznamenat, že jakákoli diferencovatelná funkce je ve svém stacionárním bodě lokálně kvazikonvevní.

Věta 2.27 Necht' $f : H \rightarrow R$ je diferencovatelná funkce na neprázdné konvevní množině $H \subset R^n$. Potom je funkce f kvazikonvevní na množině H tehdy a jen tehdy, když je f lokálně kvazikonvevní v každém bodě $\bar{\mathbf{x}} \in H$ (vzhledem k H).

Důkaz viz [4, strana 108].

Věta 2.28 *Nechť $f : H \rightarrow R$ je diferencovatelná funkce na neprázdné konvexní množině $H \subset R^n$. Potom je funkce f konvexní na množině H tehdy a jen tehdy, když je f lokálně konvexní v každém bodě $\bar{\mathbf{x}} \in H$ (vzhledem k H).*

Důkaz viz [4, strana 111].

Lemma 2.29 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$ diferencovatelná a lokálně konvexní v bodě $\bar{\mathbf{x}} \in H$, potom je lokálně kvazikonvexní v bodě $\bar{\mathbf{x}}$. Obrácená implikace neplatí.*

Důkaz viz [4, strana 111].

Příklad 2.30 *Mějme funkci $f(x) = x^3$ na množině $H = [-1, 1]$. Funkce f je v bodě $\bar{x} = -1/2$ lokálně kvazikonvexní vzhledem k $[-1, 1]$, neboť pro všechna $x \in [-1, 1]$ splňuje vztah (2.5). Není v tomto bodě lokálně konvexní vzhledem k $[-1, 1]$, stačí ve vztahu (2.6) položit $x = -1$.*

2.2 Pseudokonvexní funkce

Dostáváme se k dalšímu zobecnění konvexních funkcí. Stejně jako u lokálně (kvazi)konvexních funkcí je třeba předpokládat diferencovatelnost funkce, za to ale nebudeme požadovat konvexitu množiny, na které je daná funkce definována.

Pseudokonvexní funkce mají v optimalizaci velké využití, neboť každý jejich stacionární bod je bodem globálního minima, což nám v některých optimalizačních úlohách jako předpoklad stačí – nemusíme tedy pracovat s konvexními funkcemi, stačí nám funkce pseudokonvexní.

Definice 2.31 *Nechť $f : H \rightarrow R$ je funkce na neprázdné množině $H \subset R^n$ diferencovatelná v bodě $\bar{\mathbf{x}} \in H$. Řekneme, že je funkce f pseudokonvexní v bodě $\bar{\mathbf{x}}$, jestliže je lokálně kvazikonvexní v bodě $\bar{\mathbf{x}}$ a pro všechna $\mathbf{x} \in H$ platí:*

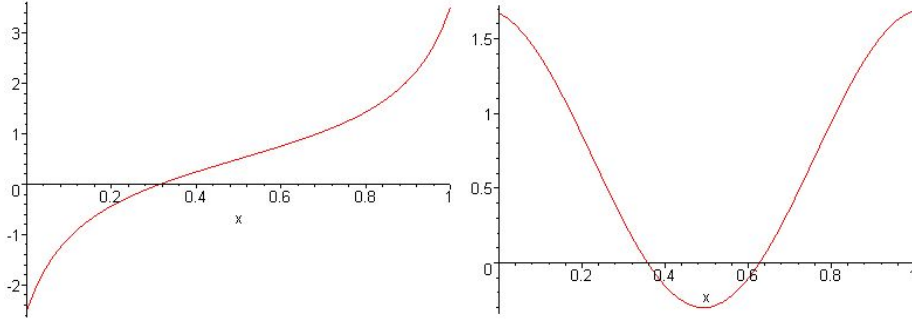
$$f(\mathbf{x}) < f(\bar{\mathbf{x}}) \Rightarrow (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) < 0 \quad (2.7)$$

Řekneme, že funkce je pseudokonvexní na otevřené množině H , jestliže je pseudokonvexní v každém bodě množiny H .

[4, strana 111].

Poznámka 2.32 *Funkce f je pseudokonkávní v bodě $\bar{\mathbf{x}} \in H$ (vzhledem k H), jestliže je funkce $-f$ pseudokonvexní v bodě $\bar{\mathbf{x}}$ (vzhledem k H).*

Nyní si uvedeme pár vztahů, které platí mezi pseudokonvexními a konvexními funkcemi a mezi pseudokonvexními a (explicitně) kvazikonvexními funkcemi. Také si uvedeme nutnou a postačující podmínku k tomu, aby byla daná funkce pseudokonvexní na konvexní množině.



Obrázek 2.4: Grafy pseudokonvexních funkcí na intervalu $[0, 1]$

Lemma 2.33 *Nechť $f : H \rightarrow R$ je funkce na neprázdné množině $H \subset R^n$. Jestliže je funkce f lokálně konvexní v bodě $\bar{\mathbf{x}} \in H$, pak je pseudokonvexní v bodě $\bar{\mathbf{x}}$. Obrácená implikace neplatí.*

Důkaz viz [4, strana 112].

Příklad 2.34 *Mějme funkci $f(x) = -(x - 1)^2$ na množině $H = [-1, 1]$. Funkce f je v bodě $\bar{x} = 0$ pseudokonvexní vzhledem k $H = [-1, 1]$, neboť splňuje vztah (2.7). Ale není v tomto bodě konvexní, stačí ve vztahu (2.6) položit $\bar{x} = 0$ a $x = -1$.*

Lemma 2.35 *Nechť $f : H \rightarrow R$ je funkce na neprázdné množině $H \subset R^n$. Jestliže je funkce f pseudokonvexní v bodě $\bar{\mathbf{x}} \in H$, pak je lokálně kvazikonvexní v bodě $\bar{\mathbf{x}}$. Obrácená implikace neplatí.*

Důkaz viz [4, strana 112].

Příklad 2.36 *Mějme funkci $f(x) = x^3$ na množině $H = [-1, 1]$. Funkce f je v bodě $\bar{x} = 0$ lokálně kvazikonvexní vzhledem k $[-1, 1]$, neboť pro všechna $x \in [-1, 1]$ splňuje vztah (2.5). Není v tomto bodě ovšem pseudokonvexní (vzhledem k $[-1, 1]$), neboť funkce f nesplňuje vztah (2.7) pro žádné $-1 \leq x < 0$.*

Věta 2.37 *Nechť $f : H \rightarrow R$ je funkce na neprázdné konvexní množině $H \subset R^n$. Potom je funkce f pseudokonvexní v bodě $\bar{\mathbf{x}} \in H$ tehdy a jen tehdy, když je diferencovatelná v bodě $\bar{\mathbf{x}}$ a implikace (2.7) platí pro všechna $\mathbf{x} \in H$.*

Důkaz viz [4, strana 113].

Věta 2.38 *Nechť $f : H \rightarrow R$ je pseudokonvexní funkce na neprázdné konvexní množině $H \subset R^n$. Potom je funkce f diferencovatelná a explicitně kvazikonvexní na H . Obrácená implikace neplatí.*

Důkaz viz [4, strana 113].

Příklad 2.39 *Mějme funkci $f(x) = x^3$ na množině $H = R$. Tato funkce je explicitně kvazikonvexní na H , ale není pseudokonvexní na H . Stačí položit ve vztahu (2.7) $\bar{x} = 0$ a za x vzít libovolné záporné číslo.*

Zde si můžeme shrnout vztahy, které platí mezi konvexními, (explicitně) kvazikonvexními a pseudokonvexními funkcemi.

Jestliže pracujeme s diferencovatelnou funkcí na neprázdné konvexní množině, pak pro danou funkci platí následující řetězec implikací:

$$\text{konvexita} \Rightarrow \text{pseudokonvexita} \Rightarrow \text{explicitní kvazikonvexita} \Rightarrow \text{kvazikonvexita}.$$

Další řetězec implikací platí pro funkci pouze v bodě $\bar{\mathbf{x}} \in H$, ve kterém je diferencovatelná (množina H nemusí být konvexní).

$$\text{konvexita} \Rightarrow \text{pseudokonvexita} \Rightarrow \text{kvazikonvexita}.$$

A poslední řetězec implikací, který zde uvedeme, platí na konvexní množině pro funkci, která nemusí být nutně diferencovatelná.

$$\text{konvexita} \Rightarrow \text{explicitní kvazikonvexita} \Rightarrow \text{kvazikonvexita}.$$

Všechny tyto řetězce najdeme například v [4, strana 114]. Podobné shrnutí vztahů mezi zobecněnými konvexními funkcemi najdeme též v [1, strana 116].

Dostáváme se k dalšímu tématu a tím jsou složené pseudokonvexní funkce. Pomocí následující věty lze rozpoznat, zda-li je daná složená funkce pseudokonvexní.

Věta 2.40 *Nechť $g : H \rightarrow R^m$ je zobrazení na neprázdné množině $H \subset R^n$, diferencovatelné a lokálně konvexní v bodě $\bar{\mathbf{x}} \in H$ vzhledem k H . Nechť $f : g(H) \rightarrow R$ je funkce na množině $g(H) \subset R^m$, pseudokonvexní v bodě $\bar{\mathbf{y}} \equiv g(\bar{\mathbf{x}})$ vzhledem k $g(H)$. Nechť $\partial f(\bar{\mathbf{y}})/\partial y_i \geq 0$ pro každé i takové, že $g_i(\mathbf{x})$ není lokálně konkávní v bodě $\bar{\mathbf{x}}$. Potom je složená funkce $\varphi(\mathbf{x}) \equiv f[g(\mathbf{x})]$ pseudokonvexní v bodě $\bar{\mathbf{x}}$ vzhledem k H .*

Důkaz viz [4, strana 114].

Předchozí věta se velmi podobá větě 2.17 a také má různé varianty, které dostaneme pokud zmodifikujeme tabulku 2.1.

Věta 2.41 *Nechť je funkce g pseudokonvexní v bodě $\bar{\mathbf{x}} \in H$ vzhledem k H a nechť je funkce f diferencovatelná v bodě $\bar{\mathbf{y}} \equiv g(\bar{\mathbf{x}})$ a $f'(\bar{\mathbf{y}}) > 0$, potom je funkce $\varphi(\mathbf{x}) \equiv f[g(\mathbf{x})]$ pseudokonvexní v bodě $\bar{\mathbf{x}}$.*

[4, strana 115].

Tato věta se velmi podobá další větě o skládání (explicitně) (kvazi)konvexních funkcí, tj. větě 2.18. Její různé varianty dostaneme, pokud zmodifikujeme tabulku 2.3.

Lemma 2.42 *Nechť $f, g : H \rightarrow R$ jsou dvě nezáporné funkce na množině $H \subset R^n$, které jsou diferencovatelné a lokálně konkávní v bodě $\bar{\mathbf{x}} \in H$ a $f(\bar{\mathbf{x}}) + g(\bar{\mathbf{x}}) > 0$. Potom je funkce $\varphi \equiv f \cdot g$ pseudokonkávní v bodě $\bar{\mathbf{x}}$ (vzhledem k množině H).*

[4, strana 116].

Jak již čtenář jistě vytušil, varianty předchozího lemmatu dostaneme, pokud zmodifikujeme tabulku 2.4.

Lemma 2.43 *Nechť $f, g : H \rightarrow R$ jsou dvě funkce na množině $H \subset R^n$, které jsou diferencovatelné v bodě $\bar{\mathbf{x}} \in H$. Nechť $f(\mathbf{x}) \geq 0$ na H a je lokálně konvexní v bodě $\bar{\mathbf{x}}$. Nechť $g(\mathbf{x}) > 0$ na H a je lokálně konkávní v bodě $\bar{\mathbf{x}}$. Potom je funkce*

$$\varphi \equiv \frac{f}{g}$$

pseudokonvexní v bodě $\bar{\mathbf{x}}$ (vzhledem k množině H).

[4, strana 116].

Varianty předešlého lemmatu dostaneme modifikací tabulky 2.5.

Lemma 2.44 *Nechť $f : H \rightarrow R$ je funkce na množině $H \subset R^n$, která je diferencovatelná a lokálně konvexní v bodě $\bar{\mathbf{x}} \in H$. Nechť $g(\mathbf{x}) \equiv \mathbf{b}^t \mathbf{x} + \beta > 0$ na množině H . Potom je funkce*

$$\varphi(\mathbf{x}) \equiv \frac{f(\mathbf{x})}{\mathbf{b}^t \mathbf{x} + \beta}$$

pseudokonvexní v bodě $\bar{\mathbf{x}}$ (vzhledem k množině H).

[4, strana 116].

Varianty tohoto lemmatu dostaneme modifikací tabulky 2.6.

2.3 Minimalizace zobecněných konvexních a konkávních funkcí

V této části si ukážeme, jak využít pěkné vlastnosti funkcí, které jsme si v předchozích částech představili. Budou to vlastnosti, kterých se využívá zejména v optimalizaci, tj. hledání minima nebo maxima těchto funkcí. Věty, které si uvedeme se budou týkat pouze minimalizace. Ovšem není těžké si uvědomit, že naprosto analogické věty platí pro maximalizaci. Při přechodu k větám o maximalizaci stačí v daných větách zaměnit (zobecněnou) konvexitu za konkávititu a naopak.

Věta 2.45 *Nechť je funkce f explicitně kvazikonvexní na konvexní množině $H \subset R^n$. Potom každý bod lokálního minima je též bod globálního minima.*

Důkaz viz [4, strana 89].

Další věta nám ukáže, za jakých předpokladů platí obrácená implikace věty předchozí.

Věta 2.46 *Nechť $f : H \rightarrow R$ je zdola polospojité funkce na konvexní množině $H \subset R^n$. Nechť pro každý uzavřený segment $S \subset H$ je každý bod lokálního minima funkce f na segmentu S též bodem globálního minima f na S . Potom je funkce f explicitně kvazikonvexní na množině H .*

Důkaz viz [4, strana 90].

Věta 2.47 *Nechť je funkce $f : R^n \rightarrow R$ pseudokonvexní. Potom $\bar{\mathbf{x}} \in R^n$ je bodem globálního minima funkce f tehdy a jen tehdy, když $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$.*

Důkaz viz [1, strana 134].

Lemma 2.48 *Nechť $f : H \rightarrow R$ je funkce na konvexní množině $H \subset R^n$. Potom je funkce f kvazikonkávní na množině H právě tehdy, když pro každý uzavřený segment $[\mathbf{x}_1, \mathbf{x}_2] \subset H$ nabývá funkce f své globální minimum v \mathbf{x}_1 nebo v \mathbf{x}_2 .*

Důkaz viz [4, strana 91].

Toto lemma bylo sice uvedeno pro minimum kvazikonkávní funkce, ale zcela analogické tvrzení platí i pro maximum kvazikonvexní funkce.

Před další větou musíme nejdříve zadefinovat globální vrcholové minimum.

Definice 2.49 *Nechť $H \subset R^n$ je konvexní polyedr. Řekneme, že vrchol $\hat{\mathbf{x}}_0 \in H$ je bodem globálního vrcholového minima funkce f na množině H , jestliže $f(\hat{\mathbf{x}}_0) \leq f(\hat{\mathbf{x}})$ pro každý vrchol $\hat{\mathbf{x}} \in H$.*

[4, strana 91].

Věta 2.50 *Nechť $f : H \rightarrow R$ je funkce na neprázdne konvexní množině $H \subset R^n$. Potom je funkce f kvazikonkávní na množině H tehdy a jen tehdy, když pro každou neprázdnu polyedrickou množinu $S \subset H$ je každý bod globálního vrcholového minima funkce f na S též bodem globálního minima f na S .*

Důkaz viz [4, strana 92].

Věta 2.51 *Nechť $f : H \rightarrow R$ je funkce na konvexním polyedru $H \subset R^n$. Jestliže je funkce f explicitně kvazikonkávní na H a nabývá svého minima na H , potom tohoto minima nabývá v nějakém vrcholu H .*

Důkaz viz [4, strana 93].

Až do teď jsme se spokojili s nalezením jednoho bodu optimálního řešení dané úlohy minimalizovat funkci na nějaké množině. Nyní se podíváme, jak vypadá množina všech bodů optimálních řešení pro zobecněné konvexní funkce.

Nejdříve si zadefinujeme právě množinu všech bodů optimálních řešení a k ní doplněk, tj. množinu všech bodů, které nejsou body optimálního řešení.

Raději ještě zdůrazníme, že se nyní bude jednat o úlohu minimalizovat danou funkci na nějaké množině.

Definice 2.52 *Nechť $f : H \rightarrow R$ je funkce na množině $H \subset R^n$. Množinu*

$$H^* \equiv \{\mathbf{x}^* \in H : f(\mathbf{x}^*) \leq f(\mathbf{x}) \forall \mathbf{x} \in H\}$$

budeme nazývat množinou optimálních řešení funkce f na množině H .

Množinu

$$\bar{H} \equiv H \setminus H^* = \{\bar{\mathbf{x}} \in H : f(\mathbf{x}) < f(\bar{\mathbf{x}}) \text{ pro nějaké } \mathbf{x} \in H\}$$

nazveme neoptimální množinou funkce f na množině H .

[4, strana 95].

Neoptimální množina je tedy množina těch přípustných řešení (tj. bodů z H), které nejsou optimální.

Nyní si uvedeme několik tvrzení, která popisují vlastnosti množiny optimálních řešení.

Lemma 2.53 *Nechť $f : H \rightarrow R$ je kvazikonvexní funkce na konvexní množině $H \subset R^n$. Potom je množina H^* konvexní.*

Důkaz viz [4, strana 95].

Lemma 2.54 *Nechť $f : H \rightarrow R$ je kvazikonkávní funkce na konvexní množině $H \subset R^n$. Potom je množina \bar{H} konvexní.*

Důkaz viz [4, strana 95].

Lemma 2.55 *Nechť $f : H \rightarrow R$ je explicitně kvazikonkávní funkce na množině $H \subset R^n$. Potom pro libovolný uzavřený segment $[\mathbf{x}_1, \mathbf{x}_2] \subset H$ je buď $[\mathbf{x}_1, \mathbf{x}_2] \subset H^*$, nebo $(\mathbf{x}_1, \mathbf{x}_2) \subset \bar{H}$.*

Důkaz viz [4, strana 96].

Lemma 2.56 *Nechť $f : H \rightarrow R$ je zdola polospojité, kvazikonvexní a explicitně kvazikonkávní funkce na konvexní polyedrické množině $H \subset R^n$. Potom H^* je konvexní obal nějakých vrcholů H .*

Důkaz viz [4, strana 96].

2.4 Některé další koncepty

Nejdříve si představíme zobecnění linearity – tzv. kvazimonotónní a explicitně kvazimonotónní funkce. Jak známo, lineární funkce je konvexní i konkávní zároveň. Kvazimonotónní funkce jsou tedy kvazikonvexní i kvazikonkávní zároveň.

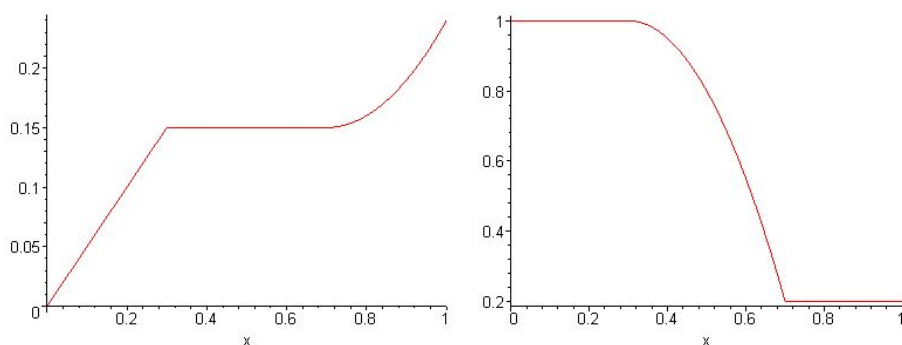
Definice 2.57 *Nechť $f : H \rightarrow R$ je funkce na konvexní množině $H \subset R^n$. Řekneme, že funkce f je kvazimonotónní na H , jestliže splňuje následující ekvivalentní podmínky:*

- (i) funkce f je kvazikonvexní i kvazikonkávní na H
- (ii) funkce f je kvazikonvexní i funkce $-f$ je kvazikonvexní na H
- (iii) pro každé dva body $\mathbf{x}_1, \mathbf{x}_2 \in H$ a $\lambda \in (0, 1)$ platí:

$$\min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\} \leq f[(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2] \leq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

- (iv) funkce f je neklesající nebo nerostoucí na každém segmentu $[\mathbf{x}_1, \mathbf{x}_2] \subset H$.

[4, strana 55].



Obrázek 2.5: Grafy kvazimonotónních funkcí na intervalu $[0, 1]$

Definice 2.58 Nechť $f : H \rightarrow R$ je funkce na konvexní množině $H \subset R^n$. Řekneme, že funkce f je explicitně kvazimonotónní na H , jestliže splňuje následující ekvivalentní podmínky:

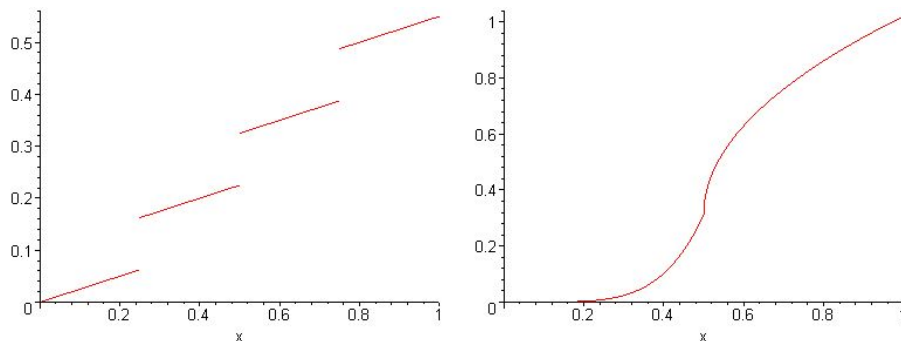
- (i) funkce f je explicitně kvazikonvexní i explicitně kvazikonkávní na H
- (ii) funkce f je explicitně kvazikonvexní i funkce $-f$ je explicitně kvazikonvexní na H
- (iii) pro každé dva body $\mathbf{x}_1, \mathbf{x}_2 \in H$ takové, že $f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$ a $\lambda \in (0, 1)$ platí:

$$\min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\} < f[(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2] < \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

- (iv) funkce f je klesající, rostoucí, nebo konstantní na každém segmentu $[\mathbf{x}_1, \mathbf{x}_2] \subset H$.

[4, strana 55].

Pro úplnost uvedeme i ekvivalentní podmínky pro linearitu funkce a po ní ještě větu, která nám říká, jak poznáme, jestli je daná funkce n proměnných lineární.



Obrázek 2.6: Grafy explicitně kvazimonotónních funkcí na intervalu $[0, 1]$

Lemma 2.59 *Nechť $f : H \rightarrow R$ je funkce na konvexní množině $H \subset R^n$. Potom je funkce f lineární na H , jestliže splňuje následující ekvivalentní podmínky:*

- (i) f je konvexní i konkávní na H
- (ii) f i $-f$ jsou konvexní na H
- (iii) pro každé dva body $\mathbf{x}_1, \mathbf{x}_2 \in H$ a $\lambda \in (0, 1)$ platí:

$$f[(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2] = (1 - \lambda)f(\mathbf{x}_1) + \lambda f(\mathbf{x}_2)$$

Důkaz viz [4, strana 56].

Věta 2.60 *Nechť $f : H \rightarrow R$ je funkce na množině $H \subset R^n$. Potom je funkce f lineární tehdy a jen tehdy, když existuje $\alpha \in R^n$ a $\beta \in R$ takové, že $f(\mathbf{x}) = \alpha^t \mathbf{x} + \beta$.*

Důkaz viz [4, strana 56].

Nyní si uvedeme definici lokální kvazimonotónie. K tomu opět nevyžadujeme konvexitu množiny, na které pracujeme. Za to ovšem potřebujeme, aby byla daná funkce diferencovatelná v bodě, kde lokální kvazimonotónii zkoumáme.

Definice 2.61 *Nechť $f : H \rightarrow R$ je funkce na množině $H \subset R^n$, která je diferencovatelná v bodě $\bar{\mathbf{x}} \in H$. Řekneme, že je funkce f lokálně kvazimonotónní v bodě $\bar{\mathbf{x}}$ (vzhledem k množině H), jestliže pro všechna $\mathbf{x} \in H$ platí:*

$$[f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) \Rightarrow (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) \leq 0] \wedge [f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) \Rightarrow (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) \geq 0]. \quad (2.8)$$

[4, strana 108].

Vztah (2.8) z předchozí definice neříká nic jiného, než že funkce kvazimonotónní v bodě $\bar{\mathbf{x}}$ musí být v bodě $\bar{\mathbf{x}}$ kvazikonvexní i kvazikonkávni. Stejně jako u lokální (kvazi)konvexity, či (kvazi)konkávity platí, že je-li funkce kvazimonotónní v každém bodě konvexní množiny, pak je na této množině kvazimonotónní.

Další věta říká, jak poznat jestli je daná diferencovatelná funkce kvazimonotónní. Zde (narozdíl od definice) už potřebujeme, aby byla množina, na které pracujeme, konvexní.

Věta 2.62 *Nechť $f : H \rightarrow R$ je diferencovatelná funkce na konvexní množině $H \subset R^n$. Potom je funkce f kvazimonotónní tehdy a jen tehdy, když pro všechna $\mathbf{x}_1, \mathbf{x}_2 \in H$ a $\mathbf{x}_0 \in [\mathbf{x}_1, \mathbf{x}_2]$ platí:*

$$f(\mathbf{x}_2) \leq f(\mathbf{x}_1) \Rightarrow (\mathbf{x}_1 - \mathbf{x}_2)^t \nabla f(\mathbf{x}_0) \leq 0.$$

Důkaz viz [4, strana 109].

Další koncept, který má ke kvazimonotónii velmi blízko, je pseudomonotónie. Nejdříve si uvedeme tedy definici pseudomonotónní funkce, a pak si ukážeme větu

Definice 2.63 *Nechť $f : H \rightarrow R$ je funkce na množině $H \subset R^n$ diferencovatelná v bodě $\bar{\mathbf{x}} \in H$. Řekneme, že je funkce f lokálně pseudomonotónní v bodě $\bar{\mathbf{x}}$ (vzhledem k množině H), jestliže je lokálně kvazimonotónní v bodě $\bar{\mathbf{x}}$ a pro všechna $\mathbf{x} \in H$ platí:*

$$[f(\mathbf{x}) < f(\bar{\mathbf{x}}) \Rightarrow (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) < 0] \wedge [f(\mathbf{x}) > f(\bar{\mathbf{x}}) \Rightarrow (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) > 0].$$

[4, strana 112].

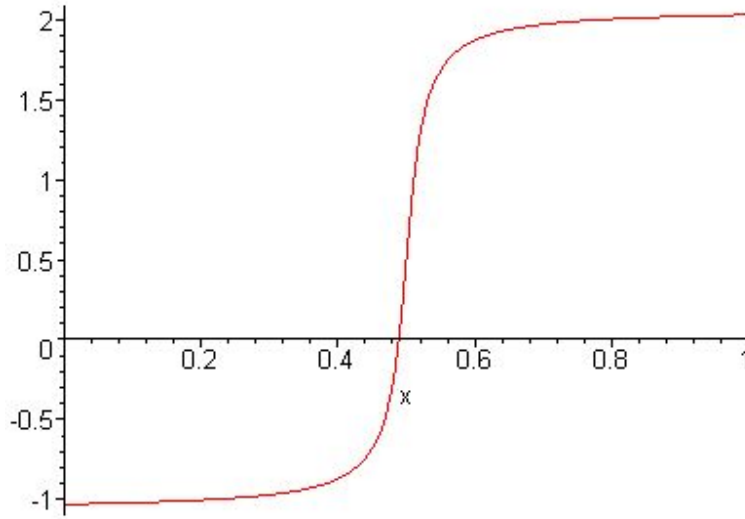
Předchozí definice opět neříká nic jiného než, že pseudomonotónní funkce je pseudokonvexní i pseudokonkávni zároveň. A opět platí, že funkce, která je pseudomonotónní v každém bodě nějaké množiny, je na této množině pseudomonotónní.

Následuje pár vět o kvazimonotónních, či pseudomonotónních funkcích. Nejdříve si ovšem pro usnadnění zápisu zavedeme značení pro dva typy množin. Pro funkci f definované na množině H si tedy zavedeme množiny:

$$\gamma(\alpha) \equiv \{\mathbf{x} \in H : f(\mathbf{x}) = \alpha\}$$

$$L_\alpha \equiv \{\mathbf{x} \in H : f(\mathbf{x}) \leq \alpha\}$$

Na začátku sekce o kvazikonvexních funkcích jsme se dozvěděli, že množina L_α je pro kvazikonvexní funkce konvexní, a tak je konvexní i pro pseudomonotónní funkce definované na konvexní množině. Nyní si ukážeme, jak tato množina vypadá pro pseudomonotónní funkce.



Obrázek 2.7: Graf pseudomonotónní funkce na intervalu $[0, 1]$

Věta 2.64 *Nechť $f : H \rightarrow R$ je pseudomonotónní funkce na konvexní množině $H \subset R^n$. Nechť $\alpha \in R$ je nějaká hodnota, které funkce f dosahuje na množině H . Potom pro libovolné $\bar{\mathbf{x}} \in \gamma(\alpha)$ platí:*

$$L_\alpha = \{\mathbf{x} \in H : (\mathbf{x} - \bar{\mathbf{x}})^t \nabla f(\bar{\mathbf{x}}) \leq 0\}.$$

Důkaz viz [4, strana 117].

V sekci o kvazikonvexních a explicitně kvazikonvexních funkcích jsme si uváděli větu 2.17 o skládání zobecněných konvexních funkcí, po které následovala různá tvrzení, která byla jejími důsledky. Následující lemma je rovněž důsledkem věty 2.17. Po tomto lemmatu bude následovat další lemma, které je důsledkem věty 2.40.

Lemma 2.65 *Lineární lomená funkce*

$$\varphi(\mathbf{x}) \equiv \frac{\mathbf{a}^t \mathbf{x} + \alpha}{\mathbf{b}^t \mathbf{x} + \beta}$$

je explicitně kvazimonotónní na konvexní množině $H \subset R^n$, jestliže $\mathbf{b}^t \mathbf{x} + \beta \neq 0$ na H .

Důkaz viz [4, strana 64].

Lemma 2.66 *Lineární lomená funkce*

$$\varphi(\mathbf{x}) \equiv \frac{\mathbf{a}^t \mathbf{x} + \alpha}{\mathbf{b}^t \mathbf{x} + \beta}$$

je pseudomonotónní v každém bodě $\bar{\mathbf{x}}$ množiny $H \subset R^n$, jestliže $\mathbf{b}^t \bar{\mathbf{x}} + \beta > 0$ na H nebo $\mathbf{b}^t \bar{\mathbf{x}} + \beta < 0$ na H .

[4, strana 116]

Lemma 2.67 *Nechť $f : H \rightarrow R$ je zdola spojitá a kvazimonotónní funkce na konvexním polyedru $H \subset R^n$. Potom množina H^* všech bodů globálního minima funkce f na konvexním polyedru H je opět konvexní polyedr.*

Důkaz viz [4, strana 96].

Na závěr zdůrazníme, že pro kvazimonotónní, explicitně kvazimonotónní, či pseudomonotónní funkce platí některé z vět uvedených v sekci – minimalizace zobecněných konvexních a konkávních funkcí. Platnost vět vyplývá z toho, že (explicitně) kvazimonotónní, respektive pseudomonotónní funkce jsou (explicitně) kvazikonvexní i kvazikonkávni, respektive pseudokonvexní i pseudokonkávni.

Tím bychom skončili naše pojednání o (explicitně)kvazimonotónních a pseudomonotónních funkcích a vrhneme se na další zajímavý koncept, který spojuje kvazikonvexní, explicitně kvazikonvexní a konvexní funkce.

Nejdřív si zadefinujeme opět méně známý pojem tzv. ω -průměr.

Definice 2.68 *Nechť $\alpha, \beta > 0$ jsou reálná čísla, $\lambda \in [0, 1]$ a $\omega \in R \setminus \{0\}$. Potom kladné reálné číslo definované vztahem*

$$\chi_\omega \equiv \chi_\omega(\alpha, \beta, \lambda) \equiv [(1 - \lambda)\alpha^\omega + \lambda\beta^\omega]^{\frac{1}{\omega}}$$

nazveme (váženým) ω -průměrem kladných reálných čísel α a β . [4, strana 65].

Zatím jsme si definovali ω -průměr pouze pro $\omega \neq 0$. Ukážeme si tedy, jak vypadá ω -průměr pro $\omega = 0$ a navíc ještě pro $\omega = \pm\infty$.

Tak tedy pro $\omega = 0$ definujeme ω -průměr jako

$$\chi_0 \equiv \lim_{\omega \rightarrow 0} \chi_\omega = \alpha^{1-\lambda} \beta^\lambda.$$

Tento vztah je také znám pod názvem geometrický průměr.

Pro $\omega = \pm\infty$ definujeme ω -průměr opět limitou a dostáváme

$$\chi_{+\infty} \equiv \lim_{\omega \rightarrow +\infty} \chi_\omega = \max\{\alpha, \beta\}$$

$$\chi_{-\infty} \equiv \lim_{\omega \rightarrow -\infty} \chi_\omega = \min\{\alpha, \beta\}$$

Následující věta shrnuje nejdůležitější vlastnosti ω -průměru.

Věta 2.69 *Nechť $\alpha, \beta > 0$ jsou reálná čísla, $\lambda \in [0, 1]$ a $\omega \in R$.*

- (i) Pokud $\alpha = \beta$, potom $\chi_\omega = \alpha$.*
- (ii) Jestliže $\alpha > \beta$, potom $\alpha > \chi_\omega > \beta$ pro jakékoli konečné ω a $\lambda \in (0, 1)$.*
- (iii) Jestliže $\omega \geq \eta$, potom $\chi_\omega > \chi_\eta$ pro jakákoli pevná α, β, λ . Nerovnost je ostrá právě tehdy, když $\alpha \neq \beta$ a $\lambda \in (0, 1)$.*

Důkaz viz [4, strana 66].

Nyní si zavedeme ω -konvexní funkce.

Definice 2.70 *Nechť $f : H \rightarrow R$ je funkce na konvexní množině $H \subset R^n$. Řekneme, že funkce f je ω -konvexní na množině H , jestliže pro všechna $\mathbf{x}, \mathbf{y} \in H$ a všechna $\lambda \in (0, 1)$ platí:*

$$f[(1 - \lambda)\mathbf{x} + \lambda\mathbf{y}] \leq \log \chi_\omega(e^{f(\mathbf{x})}, e^{f(\mathbf{y})}, \lambda). \quad (2.9)$$

Řekneme, že funkce je f ω -konkávní na H , jestliže ve vztahu (2.9) platí opačná nerovnost, nebo také jestliže $-f$ je konvexní na H .

[4, strana 66].

Pokud položíme $\omega = 0$, získáme konvexitu. Položením $\omega = +\infty$ ihned dostáváme kvazikonvexitu.

Z věty 2.69 (iii) ihned dostáváme následující tvrzení.

Důsledek 2.71 *Nechť $f : H \rightarrow R$ je ω -konvexní funkce na konvexní množině $H \subset R^n$. Potom je funkce i η -konvexní pro všechna $\eta > \omega$. Pokud navíc $f(\mathbf{x}) \neq f(\mathbf{y})$, pak v 2.9 platí ostrá nerovnost.*

[4, strana 67].

Pro $\omega < 0$ získáváme ještě silnější vlastnosti než konvexitu. Budeme se tedy zabývat jen $\omega \in [0, +\infty]$, protože tyto silnější vlastnosti (pro $\omega < 0$) nepotřebujeme. Z předchozího důsledku plyne, že je-li daná funkce ω -konvexní, kde $\omega < +\infty$, potom je explicitně kvazikonvexní.

Předchozí důsledek nám také dává spojitý přechod od konvexních funkcí ($\omega = 0$) ke kvazikonvexním ($\omega = +\infty$).

V další větě se dozvíme nutnou a postačující podmínku pro to, aby byla funkce ω -konvexní.

Věta 2.72 *Nechť $f : H \rightarrow R$ je funkce na konvexní množině $H \subset R^n$. Funkce f je ω -konvexní na H s $0 < \omega < +\infty$ tehdy a jen tehdy, když $\varphi(\mathbf{x}) \equiv e^{\omega f(\mathbf{x})}$ je konvexní na H .*

Důkaz viz [4, strana 67].

Z předchozí věty plyne, že pokud funkce f je ω -konvexní s $0 < \omega < +\infty$, pak ji lze složením s funkcí $e^{\omega x}$ převést na konvexní funkci. Toho se dá využít v optimalizaci, neboť výsledná konvexní funkce je z hlediska NLP ekvivalentní s původní ω -konvexní (ať už je to účelová funkce, nebo jedno z omezení) a má lepší vlastnosti. Už víme, že ω -konvexní funkce s $\omega < +\infty$ jsou explicitně kvazikonvexní. Vzniká tak otázka, jestli lze každou explicitně kvazikonvexní funkci převést na konvexní složením s vhodnou rostoucí funkcí (např.: $e^{\omega f}$). Odpověď je záporná, protože ne každá explicitně kvazikonvexní funkce je ω -konvexní.

Nyní si uvedeme dvě věty o skládání ω -konvexních funkcí.

Věta 2.73 *Nechť $g : H \rightarrow R^m$ je konvexní zobrazení na konvexní množině $H \subset R^n$ a $f : \text{conv}[g(H)] \rightarrow R$ je ω -konvexní funkce na konvexní obalu množiny $g(H)$, která je neklesající v každé souřadnici $i = 1, \dots, m$. Potom je funkce $\varphi(\mathbf{x}) \equiv f[g(\mathbf{x})]$ ω -konvexní na H .*

Důkaz viz [4, strana 68].

Věta 2.74 *Nechť $f, g : H \rightarrow R$ jsou dvě ω -konvexní funkce na konvexní množině $H \subset R^n$ a $\lambda \geq 0$ a $\mu \geq 0$. Potom je funkce*

$$\varphi(\mathbf{x}) \equiv \frac{1}{\omega} \log[\lambda e^{\omega f(\mathbf{x})} + \mu e^{\omega g(\mathbf{x})}]$$

ω -konvexní na H .

Důkaz viz [4, strana 68].

Tím bychom ukončili pojednání o ω -konvexních funkcích.

Závěr

Doufáme, že se čtenář dostatečně seznámil se zobecněnými konvexními funkcemi, tj. s jejich vlastnostmi a vztahy mezi jednotlivými typy.

Na úplný závěr zde zmíníme další koncepty, které buď přesahují rámec této práce, nebo jsou méně zajímavé pro praktické využití.

Prvním nezmíněným konceptem jsou tzv. silně kvazikonvexní funkce, které mají ještě silnější vlastnosti než explicitně kvazikonvexní funkce. Tento koncept najdeme například v [1, strana 112].

Druhý koncept, který zde nebyl uveden, je koncept midkonvexních funkcí, což jsou funkce pro něž platí Jensenova nerovnost, tj. vztah (1.2), ale jen pro racionální koeficienty α . Koncept o midkonvexních funkcích lze nalézt například v [5, strana 211].

Třetím nezmíněným konceptem jsou tzv. diferencovatelné (α, λ) -konkávni funkce, které jsou definovány skoro stejným způsobem jako funkce konkávni. Ovšem ve vzorci (1.1) pro konkávni funkce, tj. s opačnou nerovností, váhy na levé straně neodpovídají váhám na straně pravé. Váhy vpravo jsou funkcemi vah na straně levé a vážených bodů. Tento koncept lze nalézt například v [2, strana 52].

Posledním zmíněným konceptem jsou tzv. zcela konvexní funkce. Definičním oborem těchto funkcí je nějaký jednorozměrný interval a na tomto intervalu mají zcela konvexní funkce derivace všech řádů. Tento koncept lze nalézt například v [5, strana 233].

Literatura

- [1] Bazaraa M. S., Sherali H. D., Shetty C. M.: *Nonlinear Programming Theory and Methods - second edition*, Wiley, New York, 1993.
- [2] Cambini A., Castagnoli E., Martein L., Mazzoleni P., Schaible S.: *Generalized Convexity and Fractional Programming with Economic Applications*, Springer-Verlag, Berlin, 1990.
- [3] Lachout P.: *Teorie pravděpodobnosti*, Karolinum, Praha, 2004.
- [4] Martos B.: *Nonlinear Programming: Theory and Methods*, North-Holland, Amsterdam, 1975.
- [5] Roberts W. A., Varberg D. E.: *Convex Functions*, Academic Press, New York, 1973.
- [6] Zajíček L.: *Vybrané partie z matematické analýzy pro 1. a 2. ročník*, Matfyzpress, Praha, 2003.