MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Samir Bessisso

Zobecnění konvexních funkcí

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Petr Lachout, CSc.
Studijní program: Matematika (B1101)
Studijní obor: Obecná matematika

Praha 2020

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne ....

Podpis autora

Děkuji své rodině a Doc.RNDr. Petru Lachoutovi, CSc. za vedení práce.

Název práce: Zobecnění konvexních funkcí

Autor: Samir Bessisso

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Petr Lachout, CSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Konvexní funkce mají z pohledu matematické optimalizace řadu pěkných vlastností, jejich lokální minimum je i globálním minimem, mají konvexní dolní úrovňové množiny a jsou-li diferencovatelné, pak mají globální minimum ve stacionárním bodě. Pro hledání minima diferencovatelné konvexní funkce na konvexní množině můžeme proto efektivně využít například Karush-Kuhn-Tuckerovy podmínky nebo gradientní metody. Předpoklad konvexity funkce je ale docela restriktivní a k řadě námi využívaných vlastností konvexních funkcí ani není nutný. Tématem této bakalářské práce jsou konvexní funkce a jejich zobecnění, konkrétně kvazikonvexní a K-konvexní funkce, okrajově se zmíníme i o invexních funkcích. Práce shromažďuje poznatky o konvexních, kvazikonvexních a K-konvexních funkcích, které mohou být využity v matematické optimalizaci a ilustruje je na příkladech.

Klíčová slova: konvexní funkce, optimalizační úloha, minimum, minimální hodnota

Title: Generalization of convex functions

Author: Samir Bessisso

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Petr Lachout, CSc., Department of Probability and Mathematical Statistics

Abstract: Convex functions have range of useful properties that can be well utilized in mathematical optimization. For instance, their local minima is also global minima, they have convex lower level sets and if differentiable, their stationary point is also the point of global minima. For differentiable convex functions gradient methods and Karush-Kuhn-Tucker conditions can be effectively applied. On the other hand, the assumption of convexity is rather restrictive and not necessary for some of their desired properties. Theme of this thesis are convex functions and their generalizations, namely quasiconvex and K-convex functions, invex functions are also marginally mentioned. This thesis gathers knowledge about convex, quasiconvex and K-convex functions that can be used in mathematical optimization and ilustrates it on examples.

Keywords: convex function, optimization program, minimum, minimal value

Obsah

Úvod **2**

1 Klasická konvexita **3**
1.1 Konvexní množiny . . . 3
1.2 Konvexní funkce . . . 5
1.2.1 Operace zachovávající konvexitu funkcí . . . 7
1.2.2 Extrémy konvexních funkcí . . . 7
1.2.3 Spojitost konvexních funkcí . . . 8
1.2.4 Diferencovatelné konvexní funkce . . . 8
1.3 Motivace pro zobecnění konvexity . . . 9
1.3.1 Karush-Kuhn-Tuckerovy podmínky optimality . . . 10

2 Kvazikonvexní funkce **12**
2.1 Základní pojmy . . . 12
2.2 Extrémy kvazikonvexních funkcí . . . 13
2.3 Diferencovatelné kvazikonvexní funkce . . . 14
2.3.1 Pseudokonvexní funkce . . . 14
2.4 Operace zachovávající kvazikonvexitu . . . 15
2.5 Souvislost kvazikonvexity s úlohou maximalizace užitku . . . 15
2.6 Nested convex families . . . 18
2.6.1 Řešení kvazikonvexních problémů metodou bisekce . . . 18

3 Invexní funkce **19**

4 K-Konvexita **21**
4.1 Motivace . . . 21
4.2 K-konvexita v jedné dimenzi . . . 22
4.3 K-konvexita ve vyšších dimenzích . . . 24
4.3.1 Více druhů zboží, jeden dodavatel . . . 25
4.4 Obrázky . . . 26

Závěr **28**

Seznam použité literatury **29**

Úvod

Záměr této práce je, shromáždit teorii o konvexních funkcích a některých typech zobecněných konvexních funkcí, které mají využití v teorii optimalizace. Definice a věty jsou ilustrovány na příkladech, ale zároveň jsou některé věty a občas i definice primárně míněny jako ilustrativní příklad, pro lepší pochopení daných objektů. Konkrétně se v této práci zabývám kvazikonvexními a K-konvexními funkcemi.

První kapitola je věnována konvexním množinám a funkcím. Převážně se zde opakují definice a tvrzení, které lze znát z předmětů Matematická analýza a Úvod do optimalizace, ale jsou zde uvedena i rozšiřující tvrzení. Definice konvexní funkce je oproti těmto kurzům uvedena pro funkci, která může nabývat i nekonečných hodnot, neboť je výhodné v matematické optimalizaci a zvláště pak v konvexní optimalizaci připouštět funkce s nekonečnými hodnotami. V poslední sekci této kapitoly jsou formulovány úlohy nelineárního a konvexního programovánía uvedeny Karush-Kuhn-Tuckerovy podmínky. Hlavními zdroji pro tuto kapitolu byly učebnice Rockafellar (1997), Lachout (2011) a Dupačová a Lachout (2011).

Druhá kapitola se věnuje kvazikonvexním funkcím, které jsou motivovány zachováním pouze některých vlastností konvexních funkcí, jako jsou například konvexní dolní úrovňové množiny, unimodalita nebo globální extrém ve stacionárním bodě. Sekce 2.1 až 2.4 obsahují teorii o těchto funkcích a je v nich čerpáno hlavně z učebnic Bazaraa (2006) a Boyd (2004), sekce 2.5 je zaměřena na význam konvexních a kvazikonvexních funkcí v problému maximalizace užitku. Sekce 2.6 se věnuje [nested convex families], jejich souvislosti s kvazikonvexními funkcemi a využití při řešení kvazikonvexních problémů.

V kapitole 3 se krátce budeme věnovat invexním funkcím, které zobecňují konvexní a kvazikonvexní funkce. Pomocí obrázku 3.1 z učebnice Bazaraa (2006) je zde ukázáno, v jakém jsou dosud uvedená zobecnění vztahu. Tyto vztahy až na výjimku, že každá striktně kvazikonvexní funkce je kvazikonvexní, plynou přímo z definice a nejsou v této práci dokázány. Dále je v kapitole uvedena třída tzv. invexních funkcí prvního druhu [type I invex functions], která je nejširší třídou funkcí, pro které jsou Karush-Kuhn-Tuckerovy podmínky nutné i postačující k optimalitě řešení.

Poslední kapitola je o K-konvexních funkcích, jejichž zavedení je motivované problémem minimalizace nákladů na provoz skladu. Sekce 4.1 a 4.2 jsou věnovány K-konvexitě v jedné dimenzi a teorii čerpají z článku Scarf (2005). Sekce 4.3 rozšiřuje teorii z předchozích sekcí na vyšší dimenze a čerpá ze článku Gallego a Sethi (2005).

1. Klasická konvexita

1.1 Konvexní množiny

Definice 1 (Konvexní množina, Rockafellar (1997), str. 10)**.** *Řekneme, že množina $M \subset \mathbb{R}^k$ je konvexní, jestliže*

$$\forall x, y \in M \ \forall \lambda \in (0{,}1) : \lambda x + (1-\lambda) y \in M. \tag{1.1}$$

Součtu $\sum_{i=1}^{n} \lambda_i x_i$, kde $x_i \in \mathbb{R}^k$, $\lambda_i \geq 0$ a $\sum_{i=1}^{n} \lambda_i = 1$, se říká *konvexní kombinace* prvků $x_1, x_2, \ldots, x_n$.

Příklad: Konvexní $n-$úhelník se svým vnitřkem, tedy množina

$$N = \{x \in \mathbb{R}^2 | x = \sum_{i=1}^{n} \lambda_i a_i, \lambda_i \geq 0, \ \sum_{i=1}^{n} \lambda_i = 1\}, \tag{1.2}$$

kde $n \in \mathbb{N}$ a $a_1, \ldots, a_n \in \mathbb{R}^2$ jsou (konvexně nezávislé) vrcholy, je konvexní: Volme x, $y \in N$ a $\lambda \in (0{,}1)$. Protože x, $y \in N$, existují koeficienty $\lambda_1^x, \ldots, \lambda_n^x$, $\lambda_1^y, \ldots, \lambda_n^y$ takové, že

$$\lambda x + (1-\lambda) y = \lambda \sum_{i=1}^{n} \lambda_i^x a_i + (1-\lambda) \sum_{i=1}^{n} \lambda_i^y a_i = \sum_{i=1}^{n} (\lambda \lambda_i^x + (1-\lambda) \lambda_i^y) a_i.$$

Dále platí

$$\sum_{i=1}^{n} (\lambda \lambda_i^x + (1-\lambda) \lambda_i^y) = \lambda + (1-\lambda) = 1 \ \& \ \forall i \in \{1, \ldots, n\} : (\lambda \lambda_i^x + (1-\lambda) \lambda_i^y) \geq 0,$$

tedy $\lambda x + (1-\lambda) y$ náleží do N.

Definice 2 (Afinní množina, Rockafellar (1997), str. 3)**.** *Nechť M je neprázdná podmnožina $\mathbb{R}^k$. Řekneme, že množina M je afinní, jestliže*

$$\forall x, \ y \in M \ \forall \lambda \in \mathbb{R} : \lambda x + (1-\lambda) y \in M.$$

Součtu $\sum_{i=1}^{n} \lambda_i x_i$, $x_i \in \mathbb{R}^k$, $\lambda_i \in \mathbb{R}$, $\sum_{i=1}^{n} \lambda_i = 1$ se říká *afinní kombinace* prvků $x_1, x_2, \ldots, x_n$. Afinní množiny jsou konvexní, neboť splňují podmínku (1.1) pro libovolné $\lambda \in \mathbb{R}$.

Lemma 1 (Rockafellar (1997), Theorem 1.4)**.** *Nechť $b \in \mathbb{R}^m$ a A je matice typu $m \times k$. Množina $M \subset \mathbb{R}^k$ je afinní právě tehdy, když ji lze zapsat ve tvaru*

$$M = \{x \in \mathbb{R}^k | Ax = b\}.$$

Definice 3 (Rockafellar (1997), str. 10)**.** *Nechť $b \in \mathbb{R}^k$, $b \neq \mathbf{0}$ a $\beta \in \mathbb{R}$. Pak definujeme následující množiny:*

- $\{x \in \mathbb{R}^k | x^T b = \beta\}$ *a nazýváme ji nadrovinou.*
- $\{x \in \mathbb{R}^k | x^T b \leq \beta\}$ *a nazýváme ji uzavřeným poloprostorem.*
- $\{x \in \mathbb{R}^k | x^T b < \beta\}$ *a nazýváme ji otevřeným poloprostorem.*

Nadrovina s $\beta = 0$ je lineární podprostor $\mathbb{R}^k$ dimenze $k-1$, jehož báze je tvořena vektory kolmými na vektor b. Nadroviny s $\beta \neq 0$ jsou afinní podprostory $\mathbb{R}^k$ dimenze $k-1$.

Definice 4 (Lachout (2011), Definice 1.1)**.** *Pro neprázdnou množinu $S \subset \mathbb{R}^k$ definujeme:*
Afinní obal

$$\mathrm{Aff}(S) = \Big\{ \sum_{s \in I} \lambda(s)s \mid \lambda(s) \in \mathbb{R}\ \forall s \in I,\ \sum_{s \in I} \lambda(s) = 1,\ I \subset S \text{ konečná} \Big\}.$$

Nezáporný obal

$$\mathrm{pos}(S) = \Big\{ \sum_{s \in I} \lambda(s)s \mid \lambda(s) \geq 0\ \forall s \in I,\ I \subset S \text{ konečná} \Big\}.$$

Konvexní obal

$$\mathrm{conv}(S) = \Big\{ \sum_{s \in I} \lambda(s)s \mid \lambda(s) \geq 0\ \forall s \in I,\ \sum_{s \in I} \lambda(s) = 1,\ I \subset S \text{ konečná} \Big\}.$$

Pro prázdnou množinu definici rozšiřujeme následovně

$$\mathrm{pos}(\emptyset) = \{\mathbf{0}\}\ ,\ \mathrm{Aff}(\emptyset) = \mathrm{conv}(\emptyset) = \emptyset$$

Množina N z (1.2) je konvexní obal množiny bodů $\{a_1, \ldots, a_n\}$. $\mathrm{Conv}(S)$ je nejmenší konvexní množina, která obsahuje S. $\mathrm{Aff}(S)$ je nejmenší lineál (afinní podprostor) obsahující S.

Lemma 2 (Rockafellar (1997), Theorem 2.1)**.** *Průnik konvexních množin je konvexní množina.*

Definice 5 (Konvexní polyedrická množina, Lachout (2011), Definice 1.18)**.** *Množina $A \subset \mathbb{R}^k$ se nazývá konvexní polyedrická množina, jestliže je průnikem konečně mnoha uzavřených poloprostorů v $\mathbb{R}^k$.*

Definice 6 (Konvexní polyedr, Lachout (2011), Definice 1.18)**.** *Množina $A \subset \mathbb{R}^k$ se nazývá konvexní polyedr, existuje-li konečná množina $S \subset \mathbb{R}^k$ tak, že $A = \mathrm{conv}(S)$.*

Konvexní polyedry jsou omezené a uzavřené, tedy kompaktní v $\mathbb{R}^k$. Konvexní polyedrické množiny jsou vždy tvaru $\{x \in \mathbb{R}^k | x^T b_i \leq \beta_i,\ i = 1, \ldots, n\} = \{x \in \mathbb{R}^k | Ax \leq \beta\}$, kde $b_i \in \mathbb{R}^k$, A je matice typu $n \times k$, jejíž i-tý řádek je tvořen vektorem b_i^T a $\beta = (\beta_1, \ldots, \beta_n)^T$. Jedná se o množinu bodů, které řeší soustavu příslušných lineárních nerovnic.

Definice 7 (Směr množiny, Lachout (2011), Definice 1.26)**.** *Když A je neprázdná podmnožina $\mathbb{R}^k$, pak řekneme, že $s \in \mathbb{R}^k$ je směrem množiny A, jestliže existuje takový bod $a \in A$, že pro každé $\alpha > 0$ je $a + \alpha s \in A$. Množinu všech směrů A budeme označovat* $\mathsf{direct}(A)$*. Pro prázdnou množinu klademe* $\mathsf{direct}(\emptyset) = \{0\}$

Definice 8 (Krajní bod množiny, Lachout (2011), Definice 1.33)**.** *Nechť $A \subset \mathbb{R}^k$ je neprázdná konvexní množina. Řekneme, že bod $a \in A$ je krajním bodem množiny A, jestliže neexistují body x, $y \in A$, $x \neq y$ a $0 < \lambda < 1$ takové, aby $a = \lambda x + (1-\lambda)y$. Množinu všech krajních bodů množiny A budeme označovat* $\mathrm{ext}(A)$*.*

Lemma 3 (Lachout (2011), Věta 1.35)**.** *Konvexní polyedr má konečný počet krajních bodů a je jejich konvexním obalem.*

Věta 4 (Rockafellar (1997), Section 3)**.** *Nechť $C_1 \subset \mathbb{R}^n$, $C_2 \subset \mathbb{R}^m$ jsou konvexní množiny, potom jsou i následující množiny konvexní:*

- $\lambda C_1 = \{\lambda x | x \in C_1\}$, $\lambda \in \mathbb{R}$,
- *je-li* $m = n$, $C_1 + C_2 = \{x_1 + x_2 | x_1 \in C_1, x_2 \in C_2\}$,
- *je-li* $A : \mathbb{R}^n \to \mathbb{R}^k$ *lineární zobrazení,* $AC_1 = \{Ax | x \in C_1\}$,
- $C_1 \oplus C_2 = \{(x, y) \in \mathbb{R}^{n+m} | x \in C_1,\ y \in C_2\}$.

1.2 Konvexní funkce

Definice 9 (Konvexní funkce, Rockafellar (1997), str. 23)**.** *Nechť $f : \mathbb{R}^k \to \mathbb{R}^*$ je zobecněná reálná funkce. Množina*

$$\mathsf{epi}(f) = \{(x,y) \in \mathbb{R}^{k+1} : f(x) \leq y\}$$

se nazývá epigraf funkce f. Funkce f je konvexní, je-li množina $\mathsf{epi}(f)$ *konvexní v* $\mathbb{R}^{k+1}$*. Funkce f je konkávní, je-li funkce* $-f$ *konvexní.*
Doménou funkce f nazýváme množinu

$$\mathsf{Dom}(f) = \{x \in \mathbb{R}^k | f(x) < \infty\}.$$

Příklad: Funkce $|x|, x^2, e^x$, jsou konvexní. Funkce $-|x|, \log(x), x^3 \mathbb{I}_{x \leq 0}$ jsou konkávní. Dále například

$$g(x) = \begin{cases} 0, & x \in (-1,1), \\ \infty, & x \notin (-1,1). \end{cases} \tag{1.3}$$

je konvexní.

Definice 10 (Rockafellar (1997), str. 24)**.** *Řekneme, že funkce $f : \mathbb{R}^k \to \mathbb{R}^*$ je vlastní, jestliže*

$$\forall x \in \mathbb{R}^k : f(x) > -\infty \ a \ \exists\, x \in \mathbb{R}^n : f(x) < \infty,$$

tj. $\mathsf{epi}(f)$ *je neprázdná množina, která neobsahuje svislé přímky.*

Je-li funkce $f : S \to \mathbb{R}^*$ definována pouze na nějaké $S \subsetneq \mathbb{R}^k$, můžeme ji dodefinovat na celé $\mathbb{R}^k$ následovně

$$\widetilde{f}(x) = \begin{cases} f(x), & x \in S, \\ \infty, & x \notin S. \end{cases} \tag{1.4}$$

a její konvexita zůstane zachována. Z toho například plyne, že není-li $\mathsf{Dom}(f)$ konvexní množina, pak ani f není konvexní funkce. Nadále se proto můžeme zabývat pouze funkcemi, které jsou (do)definovány na celém $\mathbb{R}^k$ a nebude-li řečeno jinak, budeme vždy uvažovat funkce z $\mathbb{R}^k$ do $\mathbb{R}^*$.

Obdoba veškerých zde uvedených tvrzení o konvexních funkcích platí i pro konkávní funkce.

Věta 5 (Lachout (2011), Věta 2.7)**.** *Funkce f je konvexní právě tehdy, když* $\mathsf{Dom}(f)$ *je konvexní množina a pro každé x, $y \in \mathsf{Dom}(f)$ a $0 < \lambda < 1$ platí*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Definice 11 (Striktně konvexní funkce, Lachout (2011), Definice 2.14)**.** *Nechť* $\mathsf{Dom}(f)$ *je neprázdná konvexní množina. Funkce f je striktně konvexní, jestliže*

$$\forall x, y \in \mathsf{Dom}(f), x \neq y, \forall \lambda \in (0,1) : f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y).$$

Na rozdíl od konvexních funkcí, striktně konvexní funkce mají nejvýše jeden bod minima.

Definice 12 (Úrovňové množiny, Bazaraa (2006), str. 99)**.** *Pro funkci f definujeme:*

- *tzv. úrovňovou množinu* $\mathbf{L}_\alpha(f) = \{\boldsymbol{x} | f(\boldsymbol{x}) = \alpha\}$,
- *tzv. horní úrovňovou množinu* $\mathbf{L}_\alpha^+(f) = \{\boldsymbol{x} | f(\boldsymbol{x}) \geq \alpha\}$,
- *tzv. dolní úrovňovou množinu* $\mathbf{L}_\alpha^-(f) = \{\boldsymbol{x} | f(\boldsymbol{x}) \leq \alpha\}$.

Příklad: Různé úrovňové množiny funkcí $\sin(x)$ a $-(x^2+y^2)$.

$$\mathbf{L}_0^-(\sin) = \bigcup_{k\in\mathbb{Z}} [(2k-1)\pi, 2k\pi], \ \mathbf{L}_0^+(\sin) = \bigcup_{k\in\mathbb{Z}} [2k\pi, (2k+1)\pi] = \mathrm{cl}(\mathbb{R} \setminus \mathbf{L}_0^-(\sin)).$$

$$\mathbf{L}_\alpha(-(x^2+y^2)) = \begin{cases} \{(x,y) \in \mathbb{R}^2 | x^2+y^2 = -\alpha\}, & \alpha \leq 0, \\ \emptyset, & \alpha > 0. \end{cases}$$

$$\mathbf{L}_\alpha^+(-(x^2+y^2)) = \mathrm{conv}(\mathbf{L}_\alpha(-(x^2+y^2))).$$

Kde $\mathrm{cl}(M)$ značí uzávěr množiny M.

Definice 13 (Boyd (2004), sekce A.3.3)**.** *Konvexní funkce f se nazývá uzavřená, jestliže jsou všechny její dolní úrovňové množiny uzavřené.*

Příklad: Spojité konvexní funkce jsou uzavřené, funkce g z příkladu (1.3) je konvexní, ale není uzavřená.

Lemma 6 (Rockafellar (1997), Theorem 4.6.)**.** *Je-li funkce f konvexní, pak je i $\mathbf{L}_\alpha^-(f)$ konvexní pro libovolné $\alpha \in \mathbb{R}$.*

Důsledek: Úrovňové množiny afinních funkcí jsou konvexní.
Opačná implikace ve větě neplatí, brzy ale zavedeme třídu funkcí, která bude tímto vztahem definována ve formě ekvivalence.

Věta 7 (Jensenova nerovnost, Lachout (2011), Věta 2.20)**.** *Nechť $D \subset \mathbb{R}^k$ je konvexní množina a $f : D \to \mathbb{R}$ je konvexní funkce. Pak pro reálný náhodný vektor $X = (X_1, X_2, \ldots, X_k)^T$ s konečnou střední hodnotou a $P(X \in D) = 1$ platí $E[X] \in D$ a $f(E[X]) \leq E[f(X)]$.*

Lemma 8 (Důsledek Jennsenovy nerovnosti)**.** *Funkce $f : \mathbb{R}^k \to (-\infty, \infty]$ je konvexní právě tehdy, když pro $n \in \mathbb{N}$ a $x_1, \ldots, x_n \in \mathbb{R}^k$ platí*

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i), \ \textit{kde} \ \sum_{i=1}^n \lambda_i = 1 \ \textit{a} \ \lambda_i \geq 0, \forall i \in \{1, \ldots, n\}.$$

Pro f konvexní je lemma 8 speciálním případem Jensenovy nerovnosti pro diskrétní náhodnou veličinu X s rozdělením $P(X = x_i) = \lambda_i$.

1.2.1 Operace zachovávající konvexitu funkcí

Lemma 9 (Boyd (2004), Sekce 3.2.1)**.** *Buď $f : \mathbb{R} \times A \to \mathbb{R}$ konvexní funkce v proměnné x pro všechna $y \in A$ a $w(y)$ nezáporná funkce na A. Potom je*

$$g(x) = \int_A w(y) f(x,y) dy \tag{1.5}$$

konvexní funkce.

Lemma 10 (Důsledek lemmatu 9)**.** *Nechť $k \in \mathbb{N}$, $f_1, f_2, \ldots, f_k$ jsou konvexní funkce a $a_1, a_2, \ldots, a_k \geq 0$. Potom je $\sum_{i=1}^{k} a_i f_i$ konvexní funkce.*

Lemma 11 (Rockafellar (1997), Theorem 5.1.)**.** *Nechť $f : \mathbb{R}^k \to (-\infty, \infty]$ je konvexní funkce a $\phi : \mathbb{R} \to (-\infty,\infty]$ je neklesající konvexní funkce. Potom je $h(x) = \phi(f(x))$ konvexní funkce na $\mathbb{R}^k$ (klademe $\phi(\infty) = \infty$).*

Lemma 12 (Rockafellar (1997), Theorem 5.5.)**.** *Buď I indexová množina a f_i, $i \in I$ konvexní funkce. Potom je $\widetilde{f} = \sup\limits_{i \in I} f_i$ konvexní funkce.*

Lemma 13 (Rockafellar (1997), Theorem 5.4)**.** *Nechť $f_1, \ldots, f_m$ jsou vlastní konvexní funkce. Potom je*

$$f(x) = \inf\{f_1(x_1) + f_2(x_2) + \cdots + f_m(x_m) | x_i \in \mathbb{R}^n, x_1 + x_2 + \cdots + x_m = x\}.$$

konvexní.

Lemma 14 (Rockafellar (1997), Theorem 5.7.)**.** *Nechť $A : \mathbb{R}^n \to \mathbb{R}^m$ je lineární zobrazení a g konvexní funkce na $\mathbb{R}^m$. Pak je funkce*

$$(gA)(x) = g(Ax)$$

konvexní na $\mathbb{R}^n$. Je-li h konvexní funkce na $\mathbb{R}^n$, pak je

$$(Ah)(y) = \inf\{h(x) | Ax = y\}$$

konvexní na $\mathbb{R}^m$.

Lemma 15 (Gallego a Sethi (2005), Property 2.3)**.** *Nechť $g : \mathbb{R}^k \to \mathbb{R}$ je konvexní funkce a $\xi = (\xi_1, \ldots, \xi_k)$ je náhodný vektor. Nechť pro všechna $x \in \mathbb{R}^k$ platí $E|g(x - \xi)| \leq \infty$, potom je $g(x - \xi)$ konvexní.*

1.2.2 Extrémy konvexních funkcí

Věta 16 (Lachout (2011), Věta 2.15)**.** *Nechť $f : \mathbb{R}^k \to \mathbb{R}^*$ je vlastní konvexní funkce. Pak každé její lokální minimum na $\mathsf{Dom}(f)$ je také jejím globálním minimem na $\mathsf{Dom}(f)$ a množina všech globálních minim je konvexní.*

Věta 17 (Lachout (2011), Věta 2.16)**.** *Nechť $f : \mathbb{R}^k \to \mathbb{R}^*$ je vlastní ryze konvexní funkce, která má na $\mathsf{Dom}(f)$ nějaké lokální minimum. Pak má tato funkce globální minimum na $\mathsf{Dom}(f)$, které je jednoznačně určeno a je jejím jediným lokálním minimem.*

1.2.3 Spojitost konvexních funkcí

Definice 14 (Relativní vnitřek, Lachout (2011), definice 1.9)**.** *Buď $S \subset \mathbb{R}^k$, pak definujeme relativní vnitřek množiny S jako*

$$\operatorname{rint}(S) = \{x \in S | \exists \epsilon > 0, \mathbb{B}(x,\epsilon) \cap \operatorname{aff}(S) \subset S\}.$$

Kde $\mathbb{B}(x,\epsilon) = \{y \in \mathbb{R}^k | d(x,y) \leq \epsilon\}$ a $d(\cdot,\cdot)$ je euklidovská metrika.

Příklad: Pro množinu N z (1.2) platí v $\mathbb{R}^n$, $n \geq 3$, $\operatorname{int}(N) = \emptyset$, $\operatorname{rint}(N) = N$. Kde $\operatorname{int}(N)$ značí vnitřek množiny N.

Věta 18 (Lachout (2011), Věta 2.9)**.** *Vlastní konvexní funkce f je spojitá na* $\operatorname{rint}(\mathsf{Dom}(f))$.

1.2.4 Diferencovatelné konvexní funkce

Lemma 19 (Ekvivalentní definice konvexity, Gallego a Sethi (2005), Theorem 2.2)**.** *Funkce $f : \mathbb{R}^k \to \mathbb{R}$ je konvexní právě tehdy, když*

$$\forall x \in \mathbb{R}^k \ \forall y \in \mathbb{R}^k_+ \ \forall h \in \mathbb{R},\ h > 0 : f(x) + \frac{f(x) - f(x - hy)}{h} \leq f(x + y),$$

kde $\mathbb{R}^k_+$ značí $[0,\infty)^k$.

Věta 20 (Lachout (2011), Věta 2.29)**.** *Nechť D je konvexní otevřená podmnožina $\mathbb{R}^k$ a $f : D \to \mathbb{R}$ je funkce. Má-li f spojité parciální derivace na D, pak je f konvexní právě tehdy, když*

$$\forall x, y \in D : f(y) \geq f(x) + \nabla f(x)^T (y - x). \tag{1.6}$$

Funkce f je za předpokladů věty 20 striktně konvexní právě tehdy, když vztah (1.6) platí s ostrou nerovností pro všechna $x, y \in D,\ x \neq y$.

Lemma 21 (Důsledek věty 20)**.** *Nechť D je konvexní otevřená podmnožina $\mathbb{R}^k$ a $f : D \to \mathbb{R}$ je konvexní funkce. Pak je $x \in D$ bodem globálního minima funkce f právě tehdy, když je x stacionárním bodem f, tedy právě tehdy, když $\nabla f(x) = \boldsymbol{0}$.*

Věta 22 (Lachout (2011), Věta 2.24)**.** *Buď f funkce. Nechť $Dom(f)$ je otevřená konvexní množina, na které má f spojité druhé parciální derivace. Pak f je konvexní právě tehdy, když je matice*

$$\mathbf{H}(f(x)) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i=1,j=1}^{n,n}$$

pozitivně semidefinitní pro všechna $x \in \mathsf{Dom}(f)$.

Lemma 23 (Barto a Tůma (2019), Tvrzení 10.19)**.** *Symetrická matice A je pozitivně definitní (resp. semidefinitní) právě tehdy, když jsou všechna její vlastní čísla kladná (resp. nezáporná).*

Příklad: Ověříme konvexitu funkce $f(x,y,z) = -2x^2 - 3y^2 - 2z^2 + 8xy + 3xz + 4yz$,

$$\mathbf{H}(f(x)) = \begin{pmatrix} -4 & 8 & 3 \\ 8 & -6 & 4 \\ 3 & 4 & -4 \end{pmatrix}.$$

Označme A = $\mathbf{H}(f(x))$. Charakteristický polynom

$$p_A(\lambda) = -\lambda^3 - 14\lambda^2 + 25\lambda + 470$$

má kořeny

$$\lambda_1 \doteq -13.195, \ \lambda_2 \doteq -6.384 \text{ a } \lambda_3 \doteq 5.578.$$

Dle vět 22 a 23 funkce f není konvexní, ani konkávní.

1.3 Motivace pro zobecnění konvexity

Úloha nelineárního programování (NLP) je úloha minimalizovat (resp. maximalizovat) funkci f na množině řešení soustavy rovnic a nerovností

$$\begin{aligned} g_i(x) &\leq 0, \quad i = 1, \ldots, n, \\ h_j(x) &= 0, \quad j = 1, \ldots, m. \end{aligned}$$

Funkci f budeme říkat *účelová funkce*. Funkce g_i a h_j budeme nazývat *omezení*. Množině

$$\bigcap_{i=1}^{n} \{x \mid g_i(x) \leq 0\} \cap \bigcap_{j=1}^{m} \{x \mid h_j(x) = 0\}$$

se říká *množina přípustných řešení*. Bod x^* se nazývá *optimální řešení* úlohy NLP, jestliže náleží do přípustné množiny a účelová funkce v něm nabývá svého minima (resp. maxima). Bod x^{**} se nazývá *lokální optimální řešení* úlohy NLP, jestliže náleží do přípustné množiny a účelová funkce v něm nabývá svého lokálního minima (resp. lokálního maxima). Úlohy NLP se nejčastěji zapisují ve tvaru

$$\begin{array}{lll} \min & f(x), & \\ \text{za podmínek} & g_i(x) \leq 0, & i = 1, \ldots, n, \\ & h_j(x) = 0, & j = 1, \ldots, m. \end{array}$$

Mezi řešením minimalizační a maximalizační úlohy platí vzájemně jednoznačný vztah

$$x^* = \operatorname*{argmin}_{x \in S} f(x) \ \Leftrightarrow \ x^* = \operatorname*{argmax}_{x \in S} -f(x),$$

kde $S \subset \mathbb{R}^k$ značí množinu přípustných řešení. Nadále se proto můžeme a budeme zabývat pouze minimalizačními problémy.

Jsou-li v úloze NLP funkce f, $g_1, \ldots, g_n$ konvexní a funkce $h_1, \ldots, h_m$ afinní, pak mluvíme o úloze konvexního programování (CP). Z lemmatu 6 víme, že množina přípustných řešení CP je konvexní, neboť se jedná o průnik dolních úrovňových množin konvexních funkcí a úrovňových množin afinních funkcí. Protože body mimo množinu přípustných řešení S pro nás nejsou přípustné, můžeme funkci f předefinovat stejně jako v (1.4) a řešit úlohu $\min\{\widetilde{f}(x) \mid x \in \mathbb{R}^k\}$.

Z věty 16 víme, že k nalezení optimálního řešení stačí nalézt lokální optimální řešení, CP proto můžeme efektivně řešit například pomocí gradientních metod.

Jednou z motivací k zobecnění konvexních funkcí je otázka, zda-li je konvexita účelové funkce nutná k vlastnosti "lokální minimum na konvexní množině implikuje globální minimum na konvexní množině". V následující kapitole uvidíme, že existuje širší třída funkcí majících tuto dobrou vlastnost. Zároveň s tím i rozšíříme množinu, z které můžeme brát omezení tak, aby množina přípustných řešení zůstala konvexní.

1.3.1 Karush-Kuhn-Tuckerovy podmínky optimality

Pro úlohu NLP definujeme Lagrangeovu funkci $\mathcal{L}(x,\lambda,\mu) : \mathbb{R}^k \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ předpisem $\mathcal{L}(x,\lambda,\mu) = f(x) + \sum_{i=1}^n \lambda_i g_i(x) + \sum_{j=1}^m \mu_j h_j(x)$. Řekneme, že v bodě (x^*,λ^*,μ^*) jsou splněny *Karush-Kuhn-Tuckerovy podmínky optimality* (pro úlohu minimalizace), jestliže

1. $\boldsymbol{0} = \nabla f(x^*) + \sum_{i=1}^n \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^m \mu_j^* \nabla h_j(x^*)$,
2. $g_i(x^*) \leq 0, \forall i = 1,\dots,n$ a $h_j(x^*) = 0, \forall j = 1,\dots,m$,
3. $\lambda_i^* g_i(x^*) = 0,\ i = 1,\dots,n$,
4. $\lambda_i^* \geq 0, \forall i = 1,\dots,n$.

V případě maximalizační úlohy je první podmínka tvaru

$$\boldsymbol{0} = \nabla f(x^*) - \sum_{i=1}^n \lambda_i^* \nabla g_i(x^*) - \sum_{j=1}^m \mu_j^* \nabla h_j(x^*).$$

Věta 24. *Mějme úlohu NLP. Nechť f, g_i, a h_j, $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$, jsou spojitě diferencovatelné v bodě x^*. Je-li x^* lokálním optimálním řešením, pak za jistých podmínek regularity existují Karush-Kuhn-Tuckerovy multiplikátory $\lambda^* = (\lambda_1^*,\dots,\lambda_n^*)$ a $\mu^* = (\mu_1^*,\dots,\mu_m^*)$, takové, že (x^*,λ^*,μ^*) splňuje Karush-Kuhn-Tuckerovy podmínky.*

Pro úlohu CP, kde jsou navíc g_i spojitě diferencovatelné, je splnění Karush-Kuhn-Tuckerových podmínek v bodě (x^*,λ^*,μ^*) postačující pro optimalitu x^*.

Podmínky regularity jsou například:

- Podmínka lineární nezávislosti gradientů (LICQ) - Aktivní omezení v x^* mají lineárně nezávislé gradienty.
- Slaterova podmínka (SC) - Jedná se o úlohu CP a množina přípustných řešení má neprázdný relativní vnitřek. Tedy existuje x^*, které splňuje $g_i(x^*) < 0$, $\forall i \in \{1,\dots,n\}$ a $h_j(x^*) = 0$ $\forall j \in \{1,\dots,m\}$.

Aktivními omezeními v bodě x^* rozumíme omezení $h_1,\dots,h_m$ a taková g_i, pro která $g_i(x^*) = 0$.

Příklad: Pomocí Karush-Kuhn-Tuckerových podmínek dokážeme AG nerovnost pro $x_1,\dots,x_n \geq 0$, tedy

$$\frac{1}{n}\sum_{j=1}^n x_j \geq \left(\prod_{j=1}^n x_j\right)^{\frac{1}{n}}. \tag{1.7}$$

Uvažujme následující úlohu

$$\begin{array}{ll} \min & \sum_{j=1}^n x_j, \\ \text{za podmínek} & \prod_{j=1}^n x_j - 1 = 0, \\ & -x_1, \dots, -x_n \leq 0. \end{array}$$

Bude-li hodnota účelové funkce v optimálním řešení větší rovna n, tak nerovnost (1.7) bude platit pro $x_1, \dots, x_n$, která splňují $\prod_{j=1}^n x_j = 1$.

Lagrangeova funkce je tvaru

$$\mathcal{L}(x,\lambda,\mu) = \sum_{j=1}^n x_j - \sum_{j=1}^n \lambda_j x_j + \mu \left(\prod_{j=1}^n x_j - 1 \right).$$

Karush-Kuhn-Tuckerovy podmínky pro tuto úlohu jsou

$$\forall i \in \{1, \dots, n\} : 0 = 1 - \lambda_i + \mu \frac{1}{x_i} \prod_{j=1}^n x_j, \tag{1.8}$$

$$\prod_{j=1}^n x_j = 1 \ \& \ \forall j \in \{1, \dots, n\} : x_j \geq 0, \tag{1.9}$$

$$\forall j \in \{1, \dots, n\} : \lambda_j x_j = 0 \ \& \ \lambda_j \geq 0. \tag{1.10}$$

Z podmínek (1.10) a (1.9) plyne, že pro $j \in \{1, \dots, n\}$ je $\lambda_j = 0$. Potom využitím (1.8) a (1.9) musí platit

$$\forall i \in \{1, \dots, n\} : -x_i = \mu \prod_{j=1}^n x_j = \mu,$$

tedy $-\mu = x_1 = \dots = x_n = 1$. Poslední rovnost plyne z (1.9). Jediné aktivní omezení je $\prod_{j=1}^n x_j = 1$ (je-li $x_i = 0$ pro nějaké $i \in \{1, \dots, n\}$, pak zřejmě $\prod_{j=1}^n x_j \neq 1$), je proto splněna podmínka (LICQ) a protože jiného kandidáta na minimum nemáme, řešením úlohy je vektor $(1,1,\dots,1)$. V nerovnosti (1.7) máme proto, za předpokladu $\prod_{j=1}^n x_j = 1$,

$$\frac{1}{n} \sum_{j=1}^n x_j \geq \frac{1}{n} \sum_{j=1}^n x_j^* = \frac{1}{n} \sum_{j=1}^n 1 = 1 = \prod_{j=1}^n x_j = \left(\prod_{j=1}^n x_j \right)^{\frac{1}{n}}.$$

Nechť nyní $\prod_{j=1}^n x_j = c$, pro nějaké $c \in \mathbb{R}$. Pro $c = 0$ nerovnost (1.7) plyne z nezápornosti x_i, $i = 1, \dots, n$. Pro $c \neq 0$ použijeme substituci $y_j = x_j / \sqrt[n]{c}$. Protože $\prod_{j=1}^n y_j = 1$, využitím předchozích výsledků dostáváme

$$\frac{1}{n} \sum_{j=1}^n x_j = \frac{1}{n} \sum_{j=1}^n y_j \sqrt[n]{c} \geq \sqrt[n]{c} = \left(\prod_{j=1}^n x_j \right)^{\frac{1}{n}}.$$

2. Kvazikonvexní funkce

2.1 Základní pojmy

Definice 15 (Bazaraa (2006), Definition 3.5.1, 3.5.5, 3.5.8)**.** *Nechť f je funkce a $\mathsf{Dom}(f)$ je neprázdná konvexní množina. Řekneme, že funkce f kvazikonvexní, jestliže*

$$\forall x, y \in \mathsf{Dom}(f)\ \forall \lambda \in (0,1) : f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}.$$

Řekneme, že funkce f je striktně kvazikonvexní, jestliže

$$\forall x, y \in \mathsf{Dom}(f),\ f(x) \neq f(y)\ \forall \lambda \in (0,1) : f(\lambda x + (1-\lambda)y) < \max\{f(x), f(y)\}.$$

Řekneme, že funkce f je silně kvazikonvexní, jestliže

$$\forall x, y \in \mathsf{Dom}(f),\ x \neq y\ \forall \lambda \in (0,1) : f(\lambda x + (1-\lambda)y) < \max\{f(x), f(y)\}.$$

Funkce f je *kvazikonkávní* (resp. *striktně kvazikonkávní*, resp. *silně kvazikonkávní*) je-li $-f$ kvazikonvexní (resp. striktně kvazikonvexní, resp. silně kvazikonvexní). Funkce, která je kvazikonvexní i kvazikonkávní, se nazývá *kvazilineární*. Zřejmě každá konvexní funkce je i kvazikonvexní, opačná implikace neplatí. Veškerá tvrzení zde budou uvedena pouze pro kvazikonvexní funkce, ale analogie platí i pro kvazikonkávní a kvazilineární funkce.

Věta 25 (Bazaraa (2006), Theorem 3.5.2)**.** *Následující výroky jsou ekvivalentní:*

1. *Funkce $f : \mathbb{R}^k \to \mathbb{R}$ je kvazikonvexní.*
2. *$\forall \alpha \in \mathbb{R} : \mathbf{L}_\alpha^-(f)$ je konvexní množina.*

Důsledek: Je-li funkce f kvazilineární, pak je její úrovňová množina $\mathbf{L}_\alpha(f) = \mathbf{L}_\alpha^+(f) \cap \mathbf{L}_\alpha^-(f)$ konvexní pro všechna $\alpha \in \mathbb{R}$. Opačnou implikaci skoro jistě nesplňuje například funkce

$$\sum_{i=1}^{\infty} X_i \mathbb{I}_{[i,i+1)},$$

kde $X_1, X_2, \ldots$ je náhodný výběr z rozdělení se spojitou hustotou.

Příklad: Funkce $\log(x)$, $\lceil x \rceil$, x^3 jsou monotónní a proto kvazilineární.

Příklad (Boyd (2004), Example 3.31): Funkce $f(x_1,x_2) = x_1 x_2$ není pro $\mathsf{Dom}(f) = \mathbb{R}_+^2$ konvexní ani konkávní, plyne to z věty 22, neboť matice druhých derivací

$$\mathbf{H}(f) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

má charakteristický polynom $p(\lambda) = \lambda^2 - 1$. Její vlastní čísla jsou -1 a 1 a je tedy indefinitní. Funkce f je ale kvazikonkávní na $\mathbb{R}_+^2$, protože množina $\boldsymbol{L}_\alpha^+(f) = \{(x_1, x_2) \in \mathbb{R}_+^2 | x_1 x_2 \geq \alpha\}$ je konvexní pro libovolné $\alpha \in \mathbb{R}$, f není kvazikonkávní na $\mathbb{R}^2$.

Příklad: Funkce $\mathbb{I}_{\{x=0\}}$ je striktně kvazikonvexní, ale není kvazikonvexní.

Definice 16. *Řekneme, že funkce f je zdola polospojitá v bodě $x \in \mathbb{R}^k$ jestliže*

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall y \in \mathbb{B}(x,\delta) : f(y) > f(x) - \epsilon.$$

Příklad: Distribuční funkce diskrétní náhodné veličiny je (za předpokladu, že v definici distribuční funkce požadujeme spojitost zleva) zdola polospojitá.

Lemma 26 (Bazaraa (2006), Lemma 3.5.7)**.** *Nechť S je neprázdná konvexní podmnožina $\mathbb{R}^n$ a $f : S \to \mathbb{R}$ je striktně kvazikonvexní a zdola polospojitá. Potom je f kvazikonvexní.*

2.2 Extrémy kvazikonvexních funkcí

Lemma 27 (Boyd (2004), str. 99)**.** *Funkce f s $\mathsf{Dom}(f) \subset \mathbb{R}$ je kvazikonvexní právě tehdy, když splňuje jedno z následujících:*

- *f je neklesající*
- *f je nerostoucí*
- *existuje $c \in \mathsf{Dom}(f)$ takové, že f je nerostoucí na $(-\infty,c) \cap \mathsf{Dom}(f)$ a f je neklesající na $(c,\infty) \cap \mathsf{Dom}(f)$. Je-li f spojitá, pak je bod c je bodem globálního minima f.*

Kvazikonvexním funkcím na $\mathbb{R}$ se proto často říká unimodální.

Věta 28 (Bazaraa (2006), Theorem 3.5.3)**.** *Buď S neprázdný kompaktní konvexní polyedr v $\mathbb{R}^k$ a $f : \mathbb{R}^k \to \mathbb{R}$ spojitá kvazikonvexní funkce na S. Potom má úloha $\max\limits_{x \in S} f(x)$ optimální řešení x^*, kde $x^* \in \mathrm{ext}(S)$.*

Příklad: Mějme následující úlohu kvazikonvexního programování

$$\begin{array}{ll} \max & f(x_1,x_2) = x_1 x_2 \\ \text{za podmínek} & x_1 + x_2 \leq 4 \\ & x_1 \leq 1 \\ & x_1, x_2 \geq 0 \end{array}$$

Účelová funkce je kvazikonvexní a množina přípustných řešení je kompaktní konvexní polyedr. Dle předchozí věty bude proto optimální řešení v některém z bodů (0,0), (1,0), (1,3), (4,0). Dosazením zjistíme, že f nabývá svého maxima na množině přípustných řešení v bodě (1,3).

Jak jsme mohli vidět na příkladu s horní celou částí, lokální optimální řešení kvazikonvexních funkcí nemusí být globálním optimálním řešením. Následující věta říká, že striktně kvazikonvexní funkce tuto vlastnost již mají, neboť jejich definice vylučuje existenci konstantních segmentů mimo extrém.

Věta 29 (Bazaraa (2006), Theorem 3.5.6)**.** *Buď f striktně kvazikonvexní funkce a S neprázdná konvexní podmnožina $\mathbb{R}^k$. Je-li x^* bodem lokálního minima funkce f na S, pak f nabývá i svého globálního minima v bodě x^* na S.*

Nalezené globální optimální řešení ale nemusí být jednoznačné. Jednoznačnost zaručuje až předpoklad silné kvazikonvexity.

Věta 30 (Bazaraa (2006), Theorem 3.5.9)**.** *Buď f silně kvazikonvexní funkce a S neprázdná podmnožina $\mathbb{R}^k$. Je-li x^* lokálním optimálním řešením úlohy $\min\limits_{x \in S} f(x)$, pak je x^* i jednoznačným globálním optimálním řešením.*

2.3 Diferencovatelné kvazikonvexní funkce

Věta 31 (Bazaraa (2006), Theorem 3.5.4)**.** *Buď S neprázdná otevřená konvexní podmnožina $\mathbb{R}^k$ a $f : S \to \mathbb{R}$ diferencovatelná na S. Pak f je kvazikonvexní právě tehdy, když platí libovolný z lásledujících ekvivaletních výroků*

$$\begin{aligned} &\forall x_1,x_2 \in S : f(x_1) \leq f(x_2) \Rightarrow \nabla f(x_2)^T(x_1 - x_2) \leq 0, \\ &\forall x_1,x_2 \in S : \nabla f(x_2)^T(x_1 - x_2) > 0 \Rightarrow f(x_1) > f(x_2). \end{aligned} \tag{2.1}$$

Příklad: Využitím věty 31 ověříme, že funkce $f(x) = x^3$ je kvazilineární:

$$\begin{aligned} f(x_1) \leq f(x_2) \Leftrightarrow x_1 \leq x_2 \Rightarrow f'(x_2)(x_1 - x_2) \leq 0, \\ f(x_1) \geq f(x_2) \Leftrightarrow x_1 \geq x_2 \Rightarrow f'(x_2)(x_1 - x_2) \geq 0. \end{aligned}$$

Zároveň jsme zjistili, že na rozdíl od konvexních funkcí, pro kvazikonvexní funkce neplatí, že by jejich stacionární bod, byl bodem globálního minima.

Věta 32 (Boyd (2004), str. 101)**.** *Je-li f kvazikonvexní a dvakrát diferencovatelná, pak*

$$\forall x \in \mathsf{Dom}(f)\ \forall y \in \mathbb{R}^n : y^T \nabla f(x) = 0 \Rightarrow y^T \mathbf{H}(f(x))y \geq 0. \tag{2.2}$$

2.3.1 Pseudokonvexní funkce

Definice 17 (Bazaraa (2006), Definition 3.5.10)**.** *Buď S neprázdná konvexní podmnožina $\mathbb{R}^k$ a $f : S \to \mathbb{R}$ diferencovatelná funkce na S. Řekneme, že funkce f je pseudokonvexní, jestliže*

$$\forall x_1, x_2 \in S,\ \nabla f(x_1)^T(x_2 - x_1) \geq 0 \Rightarrow f(x_2) \geq f(x_1), \tag{2.3}$$

nebo ekvivalentně

$$\forall x_1, x_2 \in S,\ f(x_2) < f(x_1) \Rightarrow \nabla f(x_1)^T(x_2 - x_1) < 0.$$

Řekneme, že funkce f je striktně pseudokonvexní, jestliže

$$\forall x_1, x_2 \in S,\ x_1 \neq x_2\ \nabla f(x_1)^T(x_2 - x_1) \geq 0 \Rightarrow f(x_2) \geq f(x_1).$$

Řekneme, že funkce f je *pseudokonkávní* (resp. *striktně pseudokonkávní*), je-li $-f$ pseudokonvexní, (resp. striktně pseudokonvexní).

Bezprostředně z definice pseudokonvexity plyne, že je-li $\nabla f(x) = 0$, pak $f(y) \geq f(x)\ \forall y \in S$. Dále ze vztahu mezi větou 31 a definicí 17 plyne, že každá pseudokonvexní funkce je i kvazikonvexní.

Příklad: Hustota náhodné veličiny X s rozdělením $\mathcal{N}(\mu,\sigma^2)$ je pseudokonkávní. Derivace hustoty podle x je

$$f'_X(x) = \frac{\mu - x}{\sigma^2\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Protože $f'_X(x) > 0 \Leftrightarrow x < \mu$, máme

$$f'_X(x)(y - x) \leq 0 \Leftrightarrow y \leq x \leq \mu, \text{ nebo } \mu \leq x \leq y$$

a

$$\begin{aligned} y \leq x \leq \mu \Rightarrow f_X(y) \leq f_X(x), \\ \mu \leq x \leq y \Rightarrow f_X(y) \leq f_X(x). \end{aligned}$$

Platí tedy obdoba vztahu (2.3) pro pseudokonkání funkce.

2.4 Operace zachovávající kvazikonvexitu

Lemma 33 (Boyd (2004), str. 102)**.** *Buď I indexová množina a f_i, $i \in I$ kvazikonvexní funkce. Pak je i $g(x) = \sup_{i \in I}\{f_i(x)|i \in I\}$ kvazikonvexní funkce.*

Na rozdíl od konvexních funkcí, kvazikonvexní funkce nejsou uzavřené na kladné součty.

Příklad: Uvažujme funkci $f(x,y) = x^3 + y^3$, máme

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 \\ 3y^2 \end{pmatrix}, \mathbf{H}(f(x,y)) = \begin{bmatrix} 6x & 0 \\ 0 & 6y \end{bmatrix}.$$

Položme $z_0 = (2, -1)^T$ a $z_1 = (3, -12)$. Pak $z_1^T \nabla f(z_0) = 0$, ale $z_1^T \mathbf{H}(f(z_0)) z_1 = -756$, je tedy porušena podmínka (2.2) a funkce f proto není kvazikonvexní.

Věta 34 (Agrawal a Boyd (2020), Theorem 1)**.** *Nechť C je podmnožina $\mathbb{R}^k$, $f : C \to (-\infty,\infty]$ je kvazikonvexní zobrazení a $I = \{I_1, I_2, I_3\}$ rozdělení množiny $\{1, \ldots, k\}$ takové, že f je neklesající v argumentech indexovanými I_1 a nerostoucí v argumentech indexovanými I_2. Nechť M je podmnožina $\mathbb{R}^n$ a $g : M \to \mathbb{R}^k$ je zobrazení takové, že g_i jsou konvexní pro $i \in I_1$, konkávní pro $i \in I_2$ a afinní pro $i \in I_3$. Potom je složené zobrazení*

$$h = f \circ g$$

kvazikonvexní. Je-li navíc f konvexní, pak je i h konvexní.

2.5 Souvislost kvazikonvexity s úlohou maximalizace užitku

Uvažujme zákazníka, který má na výběr z množiny zboží, kterou označíme indexy 1,2, ..., L. Označme x vektor kvantit zboží a $X \subset \mathbb{R}_+^L$ množinu možných nákupů, nejběžněji je $X = \{x_i \geq 0 |\ i = 1, \ldots, L\}$. Na této množině zavedeme *relaci preferencí* $\succsim$. Jedná se o binární relaci na X, pomocí které porovnáváme možné alternativy $x,\ y \in X$, $x \succsim y$ čteme jako "x je alespoň stejně tak preferované jako y". Relace $\succsim$ indukuje další dvě relace na X.

Relaci *striktní preference* $\succ$, která je dána vztahem

$$x \succ y \Leftrightarrow x \succsim y \ \&\ \neg(y \succsim x),$$

$x \succ y$ čteme jako "x je preferovanější než y".

A relaci *indiference* $\sim$, která je dána vztahem

$$x \sim y \Leftrightarrow x \succsim y \ \&\ y \succsim x,$$

$x \sim y$ čteme jako "x je stejně preferované jako y".

Definice 18 (Colell (1995), Definice 3.B.1 - 3.B.4, 3.B.6)**.** *Řekneme, že relace $\succsim$ je:*

- *Racionální, je-li úplná a tranzitivní.*

- *Monotónní, jestliže*
$$x \in X, y > x \Rightarrow y \succsim x.$$

- *Lokálně nenasycená, jestliže*
$$\forall x \in X \ \forall \epsilon > 0 \ \exists y \in X, \ \|y - x\| \leq \epsilon : y \succ x. \tag{2.4}$$

- *Konvexní, jestliže*
$$\forall x \in X \text{ je množina } \{y \in X | y \succsim x\} \text{ konvexní.}$$

- *Spojitá, jestliže pro každou dvojici posloupností* $\{x_n\}_{n=1}^{\infty}$, $\{y_n\}_{n=1}^{\infty} \subset X$, *které splňují* $\lim_{n\to\infty} x_n = x$, $\lim_{n\to\infty} y_n = y$ *a* $x_n \succsim y_n$, $\forall n \in \mathbb{N}$ *je* $x \succsim y$.

Příklad: Nechť $v, k, z \in \mathbb{R}_+$, $v, k, z > 0$. Označme

$$V = \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix}, \ K = \begin{pmatrix} 0 \\ k \\ 0 \end{pmatrix}, \ Z = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix},$$

$X = \text{conv}(V,K,Z)$. Volme $x^* = (v^*,k^*,z^*)^T \in X$. Relaci $\succsim$ zavedeme na X vztahy

$$\begin{aligned} &\forall x \in X, \ x \neq x^* : x^* \succ x, \\ &\forall x, y \in X : x \succsim y \Leftrightarrow d(x,x^*) \leq d(y,x^*). \end{aligned} \tag{2.5}$$

Potom je $\succsim$ na M

1. úplná a tranzitivní,
2. monotónní, protože pro $x \in X$ je $\{y \in X | y > x\} = \emptyset$,
3. není lokálně nenasycená, protože x^* nesplňuje podmínku (2.4),
4. je konvexní, protože $\{y \in X | y \succsim x\} = \mathbb{B}(x^*, d(x,x^*)) \cap X$ je konvexní množina,
5. je spojitá.

Definice 19 (Colell (1995), Definition 1.B.2)**.** *Funkci* $u : X \to \mathbb{R}$ *nazveme užitkovou funkcí reprezentující relaci preferencí* $\succsim$, *jestliže platí*

$$u(x) \geq u(y) \Leftrightarrow x \succsim y.$$

Příklad: Užitková funkce určená relací $\succsim$ z (2.5) může být tvaru $u(x) = -d(x,x^*)$.

Požadavky na relaci $\succsim$ se jistým způsobem přenášejí na užitkovou funkci u. Je-li $\succsim$ monotónní, pak je u rostoucí. Je-li $\succsim$ konvexní, pak je u kvazikonkávní, tuto vlastnost lze snadno nahlédnout pomocí úrovňových množin funkce u: Nechť $\alpha \in \mathbb{R}$ a $\alpha' \in X$ splňuje $u(\alpha') = \inf\{x \in X | u(x) \geq \alpha\}$. Potom

$$\mathbf{L}_{\alpha}^{+}(u) = \{x \in X | u(x) \geq \alpha\} = \{x \in X | x \succsim \alpha'\},$$

přičemž $\{x \in X | x \succsim \alpha'\}$ je z konvexity $\succsim$ konvexní.

Lemma 35 (Colell (1995), Proposition 3.C.1)**.** *Je-li relace preferencí $\succsim$ racionální a spojitá na X, pak k ní existuje spojitá užitková funkce u.*

Vraťme se nyní k zákazníkovi. Označme $p = (p_1, \ldots, p_L)^T > 0$ vektor cen zboží a w zákazníkův disponibilní peněžní obnos. Nechť $X = \mathbb{R}^L_+$. Za předpokladu lineárních cen je množina cenově dostupných nákupů dána vztahem $B(w,p) = \{x \in X | p^T x \leq w\}$. Chceme najít množinu $x(p,w)$, pro kterou platí $x^* \in x(p,w) \Rightarrow x^* = \underset{x \in B(w,p)}{\operatorname{argmax}}\, u(x)$. Je-li u spojitá, pak řešení existuje.

Nechť je u spojitě diferencovatelná a $\succsim$ má vlastnosti z definice 18, pak můžeme problém maximalizace užitku formulovat jako úlohu kvazikonvexního programování a vyřešit ji pomocí Karush-Kuhn-Tuckerových podmínek. Z monotonie $\succsim$ musí optimální řešení splňovat $p^T x = w$.

$$\begin{array}{cc} \max & u(x), \\ \text{za podmínek} & p^T x - w = 0, \\ & -x_i \leq 0, \quad i = 1, \ldots, L. \end{array}$$

Lagrangeova funkce je tvaru

$$\mathcal{L}(x,\lambda,\mu) = u(x) + \sum_{j=1}^{L} \lambda_j x_j - \mu(p^T x - w).$$

Označme

$$u_i'(x) = \frac{\partial u(x)}{\partial x_i}.$$

Optimální řešení musí splňovat

$$\forall i \in \{1, \ldots, L\} : 0 = u_i'(x) + \lambda_i - \mu p_i, \tag{2.6}$$

$$p^T x - w = 0 \ \& \ \forall j \in \{1, \ldots, L\} : \lambda_j x_j = 0, \ x_j \geq 0, \ \lambda_j \geq 0. \tag{2.7}$$

Za předpokladu, že optimální řešení x^* splňuje $x_i^* > 0$ pro všechna $i \in \{1, \ldots, n\}$ je z (2.7) $\lambda_1 = \cdots = \lambda_L = 0$ (tento předpoklad závisí na tvaru funkce u a je splněn například v případě, kdy jsou $X_1, X_2, \ldots, X_L$ vzájemnými komplementy). Potom z (2.6) musí optimální řešení x^* splňovat

$$\mu = \frac{u_1'(x^*)}{p_1} = \frac{u_2'(x^*)}{p_2} = \cdots = \frac{u_L'(x^*)}{p_L} \ \& \ p^T x^* = w.$$

Již víme, že za našich předpokladů řešení existuje a že množina všech řešení je konvexní. Řešení bude jednoznačné za předpokladu, že je $\succsim$ *striktně konvexní*, tedy

$$\forall x, y \in X, y \sim x, y \neq x \, \forall \lambda \in (0,1) : \lambda y + (1-\lambda) x \succ x \ \& \ \lambda y + (1-\lambda) x \succ y.$$

Za předpokladu striktní konvexity $\succsim$ je u silně kvazikonkávní a z věty 30 je řešení jednoznačné.

2.6 Nested convex families

Definice 20 (Eppstein (2004), Nested convex families)**.** *Nechť Y je úplně uspořádaná množina a $K(\mathbb{R}^k)$ značí množinu všech kompaktních konvexních podmnožin $\mathbb{R}^k$. Pak definujeme [nested convex family] jako zobrazení $\Phi : Y \to K(\mathbb{R}^k)$, které splňuje*

1. $\forall \lambda_1,\lambda_2 \in Y,\ \lambda_1 < \lambda_2 : \Phi(\lambda_1) \subseteq \Phi(\lambda_2)$,
2. $\forall \lambda \in Y,\ \lambda = \inf\{\lambda' \in Y | \lambda' > \lambda\}\ : \Phi(\lambda) = \bigcap_{\lambda' > \lambda} \Phi(\lambda')$.

Má-li každá podmnožina Y infimum, tak pro libovolnou [nested convex family] $\Phi : Y \to K(\mathbb{R}^k)$ můžeme definovat funkci $q_\Phi : \mathbb{R}^k \to Y$ předpisem

$$q_\Phi(x) = \inf\{\lambda : x \in \Phi(\lambda)\}.$$

Nadále budeme brát $Y = \mathbb{R}^*$
Příklad: Pro $\Phi(\alpha) = \{(x,y) : x^2 + y^2 \leq \alpha\}$ máme $q_\Phi(x,y) = \inf\{\alpha | x^2 + y^2 \leq \alpha\} = x^2 + y^2$. Vidíme, že $\Phi(\alpha) = \boldsymbol{L}^-_\alpha(q_\Phi)$. Následující věta říká, že to není náhoda.

Lemma 36 (Eppstein (2004), Lemma 1.1)**.** *Pro každou [nested convex family] Φ a pro každé $\alpha \in \mathbb{R}^*$ je $\mathbf{L}^-_\alpha(q_\Phi) = \Phi(\alpha)$.*

Nechť $q : \mathbb{R}^k \to \mathbb{R}$ je kvazikonvexní funkce, jejíž dolní úrovňové množiny jsou omezené. Pak můžeme definovat zobrazení $\Phi_q(\alpha)$ předpisem

$$\Phi_q(\alpha) = \bigcap_{\alpha' > \alpha} \boldsymbol{L}^-_{\alpha'}(q).$$

Nejsou-li dolní úrovňové množiny q omezené, můžeme volit konvexní množinu $K \subset \mathbb{R}^k$ a definovat

$$\Phi_{q,K}(\alpha) = \bigcap_{\alpha' > \alpha} (\boldsymbol{L}^-_{\alpha'}(q) \cap K).$$

Lemma 37 (Eppstein (2004), Theorem 1.2)**.** *Pro libovolnou Φ je $\Phi_{q_\Phi} = \Phi$.*

Lemma 38 (Eppstein (2004), Theorem 1.3)**.** *Je-li q spojitá kvazikonvexní funkce, s omezenými dolními úrovňovými množinami, pak $q = q_{\Phi_q}$.*

2.6.1 Řešení kvazikonvexních problémů metodou bisekce

Mějme úlohu nalezení minima kvazikonvexní funkce q. Nechť víme, že se optimální řešení p^* nachází v intervalu $[\alpha,\beta]$ a nechť $\phi_t(x)$ je konvexní funkce, která splňuje $\phi_t(x) \leq 0 \Leftrightarrow x \in \Phi_q(t)$. Funkci ϕ_t lze volit například jako

$$\phi_t(x) = \begin{cases} 0, & x \in \Phi_q(t), \\ \infty, & x \notin \Phi_q(t). \end{cases}$$

Položme t $= \frac{\alpha+\beta}{2}$ a uvažujme následující problém

$$\begin{array}{ll} \text{najdi} & x, \\ \text{za podmínek} & \phi_t(x) \leq 0. \end{array}$$

Existuje-li přípustné řešení, pak se p^* nachází v intervalu $[\alpha, q(x)]$, v opačném případě se p^* nachází v intervalu $[t,\beta]$. V obou případech jsme původní interval zkrátili o polovinu. Úlohu kvazikonexního programování lze tedy řešit jako posloupnost úloh konvexního programování.

3. Invexní funkce

Definice 21 (Hanson (1981), Invexní funkce). *Řekneme, že funkce $f : \mathbb{R}^k \to \mathbb{R}$ je invexní, jestliže existuje funkce $\eta : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^k$, $\eta \not\equiv \boldsymbol{0}$, která splňuje*

$$\forall x,y \in \mathbb{R}^k : f(x) - f(y) \geq \eta(x,y)^T \nabla f(y). \tag{3.1}$$

Lemma 39 (Ben-Israel a Mond (1986), Theorem 1). *Funkce f je invexní právě tehdy, když je každý její stacionární bod bodem globálního minima.*

Definice 22 (Hanson (1981), Kvazi-invexní funkce). *Řekneme, že funkce $f : \mathbb{R}^k \to \mathbb{R}$ je kvazi-invexní, jestliže existuje funkce $\eta : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^k$, $\eta \not\equiv \boldsymbol{0}$, která splňuje*

$$f(x) - f(y) \leq 0 \Rightarrow \eta(x,y)^T f(y) \leq 0. \tag{3.2}$$

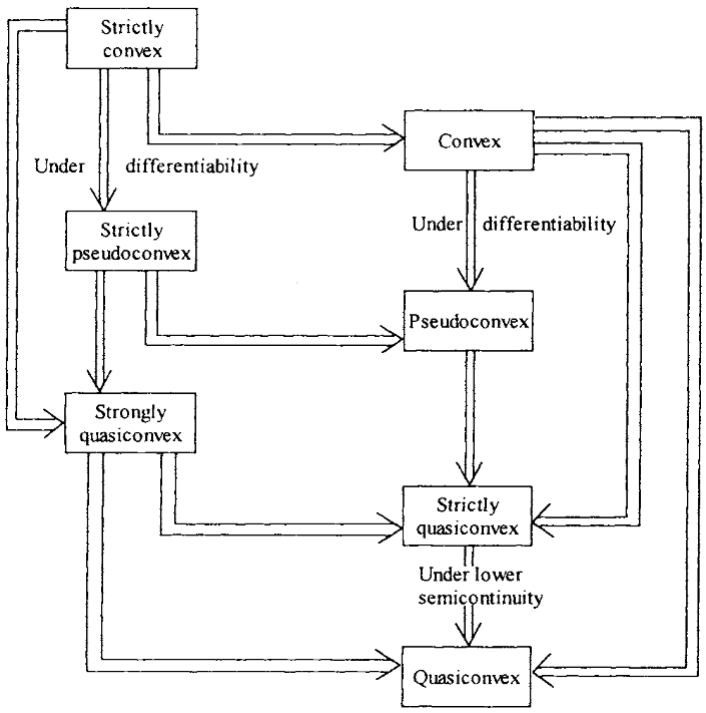
Definice 23 (Hanson (1981), Pseudoinvexní funkce). *Řekneme, že funkce $f : \mathbb{R}^k \to \mathbb{R}$ je pseudoinvexní, jestliže existuje funkce $\eta : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^k$, $\eta \not\equiv \boldsymbol{0}$, která splňuje*

$$\eta(x,y)^T \nabla f(y) \geq 0 \Rightarrow f(x) - f(y) \geq 0. \tag{3.3}$$

Podobně jako u kvazikonvexních funkcí, platí pro invexní funkce

Invexní $\Rightarrow$ Pseudoinvexní $\Rightarrow$ Kvazi-invexní.

Invexní funkce jsou rozšířením diferencovatelných konvexních funkcí, neboť volbou $\eta(x,y) = (x - y)$ dostaneme v (3.1) (resp. v (3.2), resp. v (3.3)) diferencovatelné konvexní (resp. diferencovatelné kvazikonvexní, resp. pseudokonvexní) funkce. Dále platí následující vztahy



Obrázek 3.1: Bazaraa (2006), Figure 3.13, vztahy mezi různými druhy konvexity

Příklad: Funkce $f(x_1,x_2) = x_1^3 + x_2^3$ je kvazi-invexní. Podmínku

$$x_1^3 + x_2^3 - y_1^3 - y_2^3 \leq 0 \Rightarrow 3\eta_1(x,y)y_1^2 + 3\eta_2(x,y)y_2^2 \leq 0$$

splňuje například

$$\eta(x,y) = \left(\frac{x_1^3 - y_1^3}{3y_1^2}, \frac{x_2^3 - y_2^3}{3y_2^2}\right).$$

Zobecněním invexních funkcí jsou tzv. *invexní funkce prvního druhu* [type I invex functions].

Definice 24 (Hanson (1999), type I invex function)**.** *Uvažujme úlohu*

$$\begin{array}{ll} \min & f(x), \\ \text{za podmínek} & g(x) \leq \boldsymbol{0}. \end{array} \tag{3.4}$$

Označme $F = \{x : g(x) \leq \boldsymbol{0}\}$. O funkci $f : \mathbb{R}^k \to \mathbb{R}$ řekneme, že je účelovou funkcí prvního druhu [type I objective function] v bodě $u \in F$ a o funkci $g : \mathbb{R}^k \to \mathbb{R}^m$ řekneme, že je omezením prvního druhu [type I constraint function] v bodě u, jestliže existuje funkce $\eta : F \to \mathbb{R}^k$, která splňuje

$$\begin{aligned} \forall x \in F : f(x) - f(u) &\geq \eta(x)^T \nabla f(u), \\ -g(u) &\geq \eta(x)^T \nabla g(u). \end{aligned} \tag{3.5}$$

Invexní funkce prvního druhu jsou pro nás zajímavé, protože se jedná o nejširší třídu funkcí, pro které jsou Karush-Kuhn-Tuckerovy podmínky nutné i postačující k optimalitě řešení.

Věta 40 (Hanson (1999), Theorem 2.1)**.** *Nechť $u \in S$. Jsou-li f a aktivní složky g invexní prvního druhu vzhledem ke společnému $\eta(x,u)$ v bodě u a Karush-Kuhn-Tuckerovy podmínky jsou splněny v bodě u pro nějaké λ^*, pak je u optimálním řešením problému (3.4).*

Věta 41 (Hanson (1999), Theorem 2.2)**.** *Nechť $u \in S$ a počet aktivních omezení v bodě u je menší než k (dimenze x). Je-li u optimální řešení úlohy (3.4), pak jsou funkce f a g invexní prvního druhu vzhledem ke společnému $\eta(x,u) \not\equiv \boldsymbol{0}$ v bodě u pro každé $x \in S$.*

Předpoklad o počtu aktivních omezení v bodě u není omezující, protože můžeme do (3.4) zavést proměnné $x_{k+1}, \ldots, x_{k+p}$, kde p je vhodně velké a funkci f předefinovat na

$$\widetilde{f}(x_1, \ldots, x_{k+p}) = f(x_1, \ldots, x_k) + x_{k+1}^2 + \cdots + x_{k+p}^2.$$

4. K-Konvexita

4.1 Motivace

Zavedení K-konvexity je motivované následujícím problémem: Mějme obchod, který prodává zboží X. Označme $c(x)$ cenu, za kterou obchod koupí x jednotek X. Funkce $c : \mathbb{R}_+ \to (0,\infty)$ bývá nejčastěji tvaru $c(x) = kx + K_0$, kde $k > 0$ je cena za jednotku X a $K_0 \geq 0$ je cena dovozu. Náklady na skladování zboží označíme jako $h(\cdot)$ a funkci peněžní ztráty při nedostatku zboží k prodeji jako $p(\cdot)$. Předpokládejme, že se poptávka po X za jedno prodejní období řídí nezápornou náhodnou veličinou D s hustotou ϕ. Dále předpokládejme, že zboží lze zakoupit pouze na počátku prodejního období a máme jej k dispozici okamžitě po zakoupení a že náklady jsou vyhodnoceny na konci prodejního období.

Máme-li na počátku období naskladněno y jednotek X, pak jsou očekávané náklady za toto období, dány vztahem

$$\mathrm{L}(y) = \int_0^y h(y-\xi)\phi(\xi)d\xi + \int_y^\infty p(\xi - y)\phi(\xi)d\xi. \tag{4.1}$$

Označme $\alpha \in [0,1]$ diskontní faktor, bude-li náš obchod prodávat zboží po dobu n období, očekávané náklady za jedno období při optimálním řízení a počáteční kvantitě zboží x jsou dány vztahem

$$\mathrm{C}_n(x) = \min_{y \geq x}\{c(y-x) + \mathrm{L}(y) + \alpha \int_0^\infty \mathrm{C}_{n-1}(y-\xi)\phi(\xi)d\xi\}, \tag{4.2}$$

kde $C_0 \equiv 0$. Jestliže $y_n(x)$ je takové y, které minimalizuje (4.2), pak je optimální počáteční koupě $y_n(x) - x$.

Příklad: Mějme kamelota, který se musí ráno rozhodnout, kolik novin má od svého dodavatele nakoupit, aby maximalizoval své zisky. Označme h pořizovací cenu za jedny noviny a p výdělek za jejich prodej. Funkce ztrát při nedostatku novin k prodeji je tvaru $p(x) = px$. Neprodané noviny kamelot na konci dne věnuje svému mladšímu bratrovi, který chodí na základní školu, kde mají celoroční soutěž ve sběru papíru. Funkce nákladů za skladování je proto tvaru $h(x) = hx$ a počet prodejních období je jedna. Budeme předpokládat, že D je spojitá a že má kladnou hustotu.

Kamelot koupí nějaký počet novin a ihned po jejich zakoupení potká svého kamaráda z MFF UK a vylíčí mu svou situaci. Kamarád z MFF UK problém formuluje jako úlohu matematického programování (pro názornost minimalizační)

$$\min_{y \geq 0} \mathrm{L}(y),$$

jejíž řešení je v dané situaci i řešením úlohy maximalizace zisků

$$\max_{y \geq 0} \int_0^y (p+h)x\phi(x)dx - hy.$$

Lagrangeova funkce pro minimalizační problém je tvaru $\mathcal{L}(y,\lambda) = \mathrm{L}(y) - \lambda y$, pro řešení dostáváme, využitím derivace integrálu podle meze, Karush-Kuhn-Tuckerovy podmínky

$$0 = hF_D(y) - p(1 - F_D(y)) - \lambda,\ \lambda \geq 0,\ y \geq 0,\ \lambda y = 0.$$

Jediné netriviální řešení je

$$y = F_D^{-1}\left(\frac{p}{h+p}\right), \ \lambda = 0. \tag{4.3}$$

Označme S_1 takové y, které splňuje (4.3). Protože L je lemmat 9 a 15 konvexní a množina přípustných řešení má neprázdný vnitřek, je splněna podmínka (SC) a L nabývá na $[0,\infty)$ svého minima v S_1. Dostáváme se k závěru

$$\mathrm{C}_1(x) = \begin{cases} \mathrm{L}(S_1), & x < S_1, \\ \mathrm{L}(x), & x > S_1. \end{cases}$$

Kamarád z MFF UK proto poradí kamelotovi, aby v případě, že má méně než S_1 novin doplnil své zásoby na S_1 (v praxi $\lfloor S_1 \rfloor$ nebo $\lceil S_1 \rceil$), protože při tomto počtu novin bude mít v našem modelu maximální očekávaný výdělek. V opačném případě mu řekne, aby další noviny nekupoval a popřeje mu hodně štěstí.

Nechť $s, S \in \mathbb{R}$, $s \leq S$. Řízení skladu, kdy zboží doplníme na hladinu S pouze v případě, kdy je $x < s$, a v opačném případě nic neděláme, se v teorii skladu říká (s, S) řízení.

Předchozí příklad byl na maximalizaci zisku za jedno období při nulové ceně dovozu. Dá se dokázat, že pro $K_0 = 0$ a $c(x), h(\cdot)$, $p(\cdot)$ lineární, existuje posloupnost čísel $S_1, S_2, \ldots$ taková, že je řízení (S_n,S_n) při horizontu n období optimální.

Matematici padesátých let se zabývali otázkou, jak se situace změní, když bude $K_0 > 0$. Odpověď publikoval na konci padesátých let matematik Herbert Scarf ve svém článku Scarf (2005), ve kterém zavedl třídu K-konvexních funkcí.

4.2 K-konvexita v jedné dimenzi

Definice 25 (Scarf (2005), definice K-konvexní funkce)**.** *Nechť $K \geq 0$. Řekneme, že funkce $f : \mathbb{R} \to \mathbb{R}$ je K-konvexní, jestliže*

$$\forall x \in \mathbb{R} \ \forall a, b > 0 : K + f(a+x) - f(x) - a\left(\frac{f(x) - f(x-b)}{b}\right) \geq 0.$$

Lemma 42 (Gallego a Sethi (2005), Theorem 2.2)**.** *Funkce f je K-konvexní právě tehdy, když*

$$\forall x,y \in \mathbb{R},\ y \geq x \ \forall \lambda \in [0,1] : f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)[K + f(y)].$$

Lemma 43 (Scarf (2005))**.** *Konvexní funkce jsou 0-konvexní.*

Příklad: Funkce

$$-\mathrm{sgn}(x) = \begin{cases} 1, & x < 0, \\ 0, & x = 0, \\ -1, & x > 0. \end{cases}$$

je 2-konvexní, funkci $\mathrm{sgn}(x)$ lze považovat za 2-konkávní.

Příklad: Funkce $\log(x)$ je na intervalu (m,M), $M > m > 0$, $\left(\frac{M-m}{m}\right)$-konvexní. Z konkavity a monotonie logaritmu a věty 20 plyne, že pro všechna $x,y \in (m,M)$, $y > x$ platí

$$\log(\lambda x + (1-\lambda)y) \leq \log'(x)(\lambda x + (1-\lambda)y - x) + \log(x) =$$
$$= \frac{\lambda x + (1-\lambda)y - x}{x} + \log(x) \leq \frac{\lambda x + (1-\lambda)y - x}{x} + \lambda \log(x) + (1-\lambda)\log(y) =$$
$$= \lambda \log(x) + (1-\lambda)\left[\log(y) + \frac{y-x}{x}\right] \leq \lambda \log(x) + (1-\lambda)\left[\log(y) + \frac{M-m}{m}\right].$$

Lemma 44 (Scarf (2005)). *Je-li f K-konvexní, pak je i $g(x) = f(x+h)$ K-konvexní pro všechna $h \in \mathbb{R}$.*

Lemma 45 (Scarf (2005)). *Nechť f je K-konvexní a g je M-konvexní. Potom $\alpha f + \beta g$ je $\alpha K + \beta M$ konvexní. Tuto vlastnost lze zobecnit na nespočetné sumy a integrály.*

Položme

$$\mathrm{G}_n(y) = ky + \mathrm{L}(y) + \alpha \int_0^\infty \mathrm{C}_{n-1}(y-\xi)\phi(\xi)d\xi. \tag{4.4}$$

Potom je

$$\mathrm{C}_0 \equiv 0, \ \mathrm{C}_n(x) = \min_{y \geq x}\{G_n(y) - kx\}$$

a K_0-konvexita funkce C_n závisí pouze na K_0-konvexitě funkce G_n. Z lemmat 44 a 45 K_0-konvexita funkce G_n závisí pouze na K_0-konvexitě funkce C_{n-1}. V Scarf (2005) je pak matematickou indukcí a diskuzí o poloze x dokázáno, že

$$\mathrm{C}_n(x) = \begin{cases} K_0 - kx + G_n(S_n), & x < s_n, \\ G_n(x) - kx, & x > s_n. \end{cases}$$

Kde $S_n = \operatorname{argmin} \mathrm{G}_n(x)$ a s_n splňuje $\mathrm{G}_n(s_n) = \mathrm{G}_n(S_n) + K_0$ (lze pouze pro D spojitou, pro D diskrétní je třeba volit největší s_n, které splňuje $s_n \leq S_n$ & $\mathrm{G}_n(s_n) > \mathrm{G}_n(S_n) + K_0$).

Jinými slovy, nenastane situace na obrázku 4.1 (tedy s_n a S_n jsou jednoznačně určeny) a zboží X je výhodné koupit až v době, kdy očekávané náklady přesahují minimální očekávané náklady plus cenu za dovoz.

Příklad: $h(t) = 50t$, $p(t) = 250t$, $K_0 = 400$, $k = 75$, $D = Bi(100,0.3)$, $n = 5$ a $\alpha = 0.75$. Potom pro $y \in \mathbb{N}$

$$\mathrm{L}(y) = \sum_{j=0}^{y} 50(y-j)\binom{100}{j}(0.3)^j(0.7)^{100-j} + \sum_{j=y+1}^{100} 250(j-y)\binom{100}{j}(0.3)^j(0.7)^{100-j}.$$

Dále, použitím programu R, dostáváme výsledky

i	s_i	S_i	$G_i(S_i)$	$G_i(s_i)$
1	25	31	2784.475	3214.834
2	20	25	3587.848	4043.533
3	19	24	3684.983	4208.604
4	19	24	3693.257	4204.251
5	19	24	3693.962	4204.306

a

$$\mathrm{C}_5(x) = \begin{cases} 4093.962 - 75x, & x < s_5, \\ \mathrm{L}(x) + \alpha \sum_{j=0}^{x} \mathrm{C}_4(x-j)\mathbb{P}(D=j), & x \geq s_5. \end{cases}$$

V sekci obrázky jsou grafy funkcí L (obr. 4.2), G_5 (obr. 4.3) a C_5 (obr. 4.4). V dalším budeme psát pouze G místo G_n, S místo S_n atd.

4.3 K-konvexita ve vyšších dimenzích

K-konvexita ve vyšších dimenzích je motivovaná obdobným problémem jako v předchozí sekci. Ve skladu nyní máme n různých druhů zboží $X_1, \ldots, X_n$, přičemž cena za dovoz je dána funkcí $\mathrm{K} : \mathbb{R}_+^n \to \mathbb{R}_+$, která je definovaná předpisem

$$\mathrm{K}(x) = K_0 \delta(e^T x) + \sum_{i=1}^{n} K_i \delta(x_i),$$

kde $e = (1,1,\ldots,1)^T \in \mathbb{R}^n$ a $\delta : \mathbb{R} \to \mathbb{R}$, $\delta(x) = \mathbb{I}_{\{x>0\}}$.

Tedy musíme zaplatit K_i kdykoli objednáváme X_i a K_0, kdykoli objednáváme.

Lemma 46 (Gallego a Sethi (2005), Lemma 2.1)**.** *Funkce* $\mathrm{K} : \mathbb{R}_+^n \to \mathbb{R}_+$ *splňuje trojúhelníkovou nerovnost.*

Lemma 47 (Gallego a Sethi (2005), Lemma 2.2)**.** *Funkce* $\mathrm{K} : \mathbb{R}_+^n \to \mathbb{R}_+$ *je homogenní stupně* 0.

Definice 26 (Gallego a Sethi (2005), Definice 2.1)**.** *Nechť* $g : \mathbb{R}^k \to \mathbb{R}$ *je funkce a* $\mathcal{K} = (K_0, K_1, \ldots, K_n)^T \in \mathbb{R}_+^{n+1}$. *Řekneme, že* g *je* $(K_0, K_1, \ldots, K_n)$*-konvexní, nebo prostě* $\mathcal{K}$*-konvexní, jestliže pro každé* $\lambda \in [0,1]$ *platí*

$$\forall x, y \in \mathbb{R}^k, y \geq x : g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)(g(y) + \mathrm{K}(y-x)).$$

$\mathcal{K}$-konvexita úzce souvisí s pojmem viditelnosti. Nechť $a \geq 0$, o bodu $(x, f(x))$ řekneme, že je *viditelný* z bodu $(y, f(y)+a)$, jestliže křivka $\{(\lambda x + (1-\lambda)y, f(\lambda x + (1-\lambda)y) | \lambda \in [0,1]\}$ leží pod úsečkou spojující body $(x, f(x))$ a $(y, f(y)+a)$.

Příklad: Nechť $f(x_1, x_2) = (2x_1, 4x_2^2)$, pak je (5,10,100) viditelný z (1,2,4), protože $\mathsf{epi}(f)$ je konvexní.

Příklad: Pro $g(x) = -x^2 + 100$, bod (10,0) není viditelný z bodu (−10,0).

Věta 48 (Gallego a Sethi (2005), Theorem 2.1)**.** *Funkce* g *je* $\mathcal{K}$*-konvexní právě tehdy, když* $(x, g(x))$ *je viditelný z* $(y, g(y) + \mathrm{K}(y-x))$ *pro všechna* $y \geq x$.

Lemma 49 (Gallego a Sethi (2005), Property 2.1)**.** *Nechť* $g : \mathbb{R}^n \to \mathbb{R}$ *je* $\mathcal{K}$*-konvexní. Potom je* g *i* $\mathcal{L}$*-konvexní pro libovolné* $\mathcal{L} \geq \mathcal{K}$.

Lemma 50 (Gallego a Sethi (2005), Property 2.2)**.** *Nechť* $g_1 : \mathbb{R}^n \to \mathbb{R}$ *je* $\mathcal{K}$*-konvexní a* $g_2 : \mathbb{R}^n \to \mathbb{R}$ *je* $\mathcal{L}$*-konvexní. Nechť* $\alpha, \beta \geq 0$. *Potom* $\alpha g_1 + \beta g_2$ *je* $(\alpha\mathcal{K} + \beta\mathcal{L})$*-konvexní.*

Lemma 51 (Gallego a Sethi (2005), Property 2.3)**.** *Nechť* $g : \mathbb{R}^n \to \mathbb{R}$ *je* $\mathcal{K}$*-konvexní a* $\xi = (\xi_1, \ldots, \xi_n)$ *je náhodný vektor, který pro každé* $x \in \mathbb{R}^n$ *splňuje* $E|g(x-\xi)| \leq \infty$. *Potom je* $Eg(x-\xi)$ $\mathcal{K}$*-konvexní.*

Lemma 52 (Gallego a Sethi (2005), Property 2.4)**.** *Nechť $g : \mathbb{R}^n \to \mathbb{R}$ je funkce. Volme $x, y \in \mathbb{R}^n$, $y \geq x$ a definujme*

$$f(\lambda) = g(\lambda x + (1-\lambda)y).$$

Funkce $f : \mathbb{R} \to \mathbb{R}$ je $\mathrm{K}(y-x)$-konvexní právě tehdy, když je g $\mathcal{K}$-konvexní.

Lemma 53 (Gallego a Sethi (2005), Theorem 2.3)**.** *Nechť $g : \mathbb{R}^n \to \mathbb{R}$ je $\mathcal{K}$-konvexní a $S \in \mathbb{R}^n$ je bodem globálního minima funkce g. Nechť $x \leq S$ a definujme $f(\lambda) = g(\lambda S + (1-\lambda)x)$. Nechť $\lambda_1 < 1$ splňuje $f(\lambda_1) = f(1) + \mathrm{K}(S-x)$. Potom je $f(\lambda)$ neklesající na $\lambda < \lambda_1$ a $f(\lambda) \geq f(1) + \mathrm{K}(S-x)$ pro všechna $\lambda \leq \lambda_1$.*

Věta 54 (Gallego a Sethi (2005), Theorem 3.1)**.** *Nechť $g_i : \mathbb{R} \to \mathbb{R}$ je $\mathcal{K}_i$-konvexní, $i = 1,\ldots,n$. Potom je funkce $g(x_1,\ldots,x_n) = \sum_{i=1}^n g_i(x_i) : \mathbb{R}^n \to \mathbb{R}$, $(0,K_1,\ldots,K_n)$-konvexní.*

Příklad: Označme $X = (X_1,X_2)$, $c(x) = c_1(x_1) + c_2(x_2)$, a $D = (D_1,D_2)$ poptávku po X. Budeme předpokládat, že D má nezávislé složky a že $G = G_1 + G_2 : \mathbb{R}^2 \to \mathbb{R}$ je $(0,K_1,K_2)$-konvexní. Pro $i \in \{1,2\}$ definujme $S_i = \operatorname{argmin} \mathrm{G}_i(x)$ a s_i vztahem $\mathrm{G}_i(s_i) = K_i + \mathrm{G}_i(S_i)$, $S = (S_1,S_2)$, $s = (s_1,s_2)$. Potom lze z našich znalostí o K-konvexitě v jedné dimenzi odvodit

$$\mathrm{C}(x) = \mathrm{C}_1(x) + \mathrm{C}_2(x) = \begin{cases} K_1 + K_2 - k_1x_1 - k_2x_2 + G(S), & x < s, \\ K_2 - k_1x_1 - k_2x_2 + G(x_1,S_2), & x_1 > s_1,\ x_2 < s_2, \\ K_1 - k_2x_2 - k_1x_1 + G(S_1,x_2), & x_1 < s_1,\ x_2 > s_2, \\ G(x) - k_1x_1 - k_2x_2, & x > s. \end{cases}$$

4.3.1 Více druhů zboží, jeden dodavatel

Nechť $\mathcal{K}_0 = (K_0,0,\ldots,0) \in \mathbb{R}^{n+1}$ a funkce $g : \mathbb{R}^n \to \mathbb{R}$ je $\mathcal{K}_0$-konvexní, spojitá a koercivní (tj. $\lim_{\|x\|\to\infty} g(x) = \infty$). Buď S bod globálního minima funkce g a $\delta\sigma = \{x \leq S | g(x) = g(S) + K_0\}$. Množina $\delta\sigma$ je díky koercivitě a spojitosti g neprázdná a omezená. Definujme

$$\Sigma = \{x \leq S | \exists s \in \delta\sigma\ \exists \lambda \in [0,1] : x = \lambda s + (1-\lambda)S\},$$
$$\sigma = \{x \leq S | x \notin \Sigma\}.$$

Lemma 55 (Gallego a Sethi (2005), Lemma 5.1)**.** *Je-li funkce $g : \mathbb{R}^n \to \mathbb{R}$ $\mathcal{K}_0$-konvexní, spojitá a koercivní, pak*

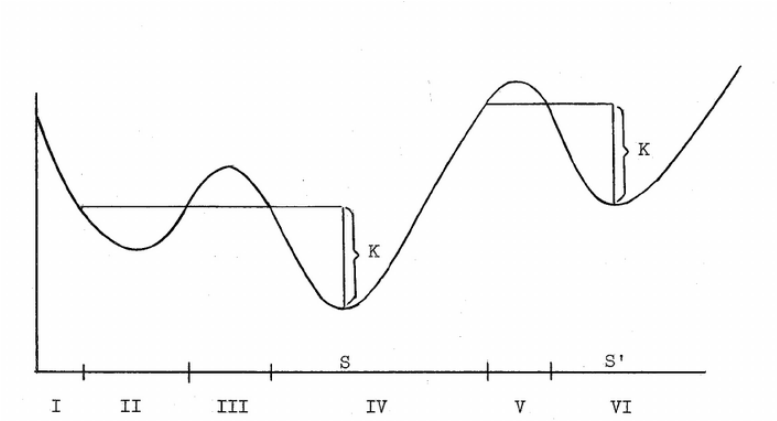
1. *$x \in \Sigma \Rightarrow g(x) \leq g(S) + K_0$,*
2. *$x \in \sigma \Rightarrow g(x) > g(S) + K_0$.*

Lemma 56 (Gallego a Sethi (2005), Lemma 5.2)**.** *Nechť $g : \mathbb{R}^n \to \mathbb{R}$ je spojitá, koercivní a $\mathcal{K}_0$-konvexní a S je bod globálního minima funkce g. Potom*

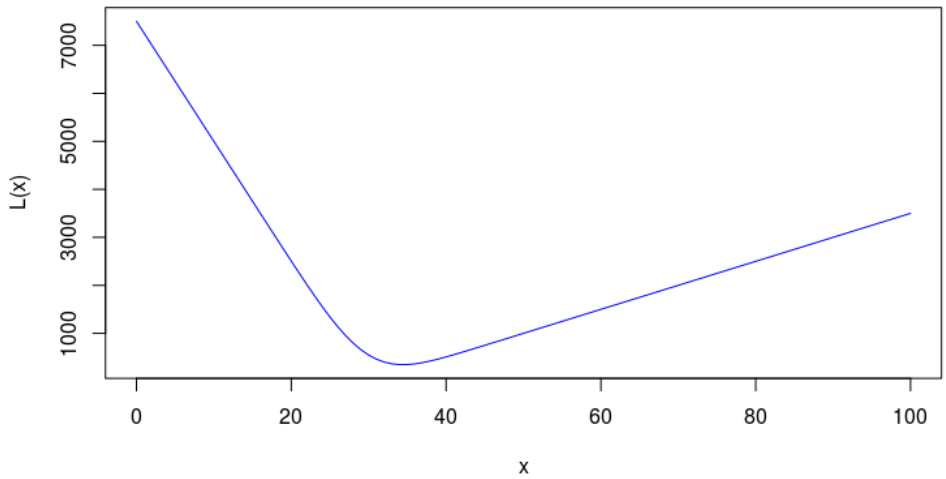
$$C(x) = \inf_{y \geq x} \{g(y) + K_0\delta(e^T(y-x))\} = \begin{cases} K_0 + g(S), & x \in \sigma, \\ g(x), & x \in \Sigma, \end{cases}$$

a funkce C je spojitá a $\mathcal{K}_0$-konvexní na $\{x \in \mathbb{R}^n | x \leq S\}$.

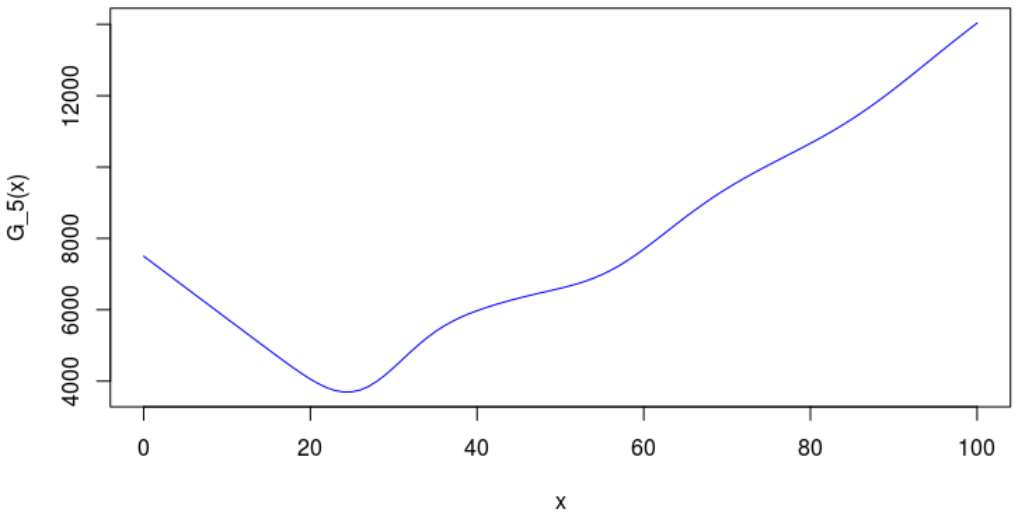
4.4 Obrázky



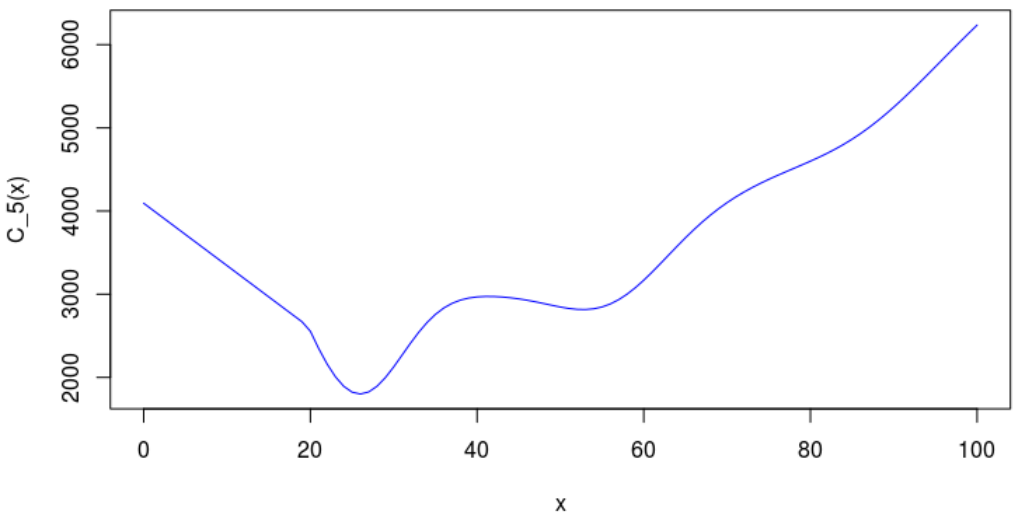
Obrázek 4.1: (Zdroj: Scarf (2005)) K-konvexita funkce G_n zaručuje, že nenastane situace, kde by jsme měli dva různé body S, S' a dva různé body s, s', které splňují $G_n(s) = G_n(S) + K$ a $G_n(s') = G_n(S') + K$. Jinými slovy funkce G_n může oscilovat, ale oscilace nikdy nebude tak velká, aby způsobila vychýlení od (s,S) řízení.



Obrázek 4.2: L



Obrázek 4.3: $G_5(x)$



Obrázek 4.4: $C_5(x)$

Závěr

U konvexních funkcí jsme uvedli jejich důležité vlastnosti pro teorii optimalizace. Tvrzení ze sekcí 1.2.1 a 2.4 mají v současné době uplatnění při řešení (kvazi)konvexních problémů za pomocí softwaru CVXPY, který umí od letošního roku řešit i kvazikonvexní problémy (Agrawal a Boyd (2020)).

U kvazikonvexních funkcí jsme ukázali, že musí být "monotónní"v tom smyslu, že jejich dolní úrovňové množiny tvoří [nested convex family]. Zavedli jsme pseudokonvexní funkce, jejichž stacionární bod je i bodem globálního minima. Dále jsme ukázali, že užitková funkce je při konvexních preferencích kvazikonkávní a následně jsme za jistých předpokladů ukázali řešení problému maximalizace užitku.

U K-konvexity v jedné dimenzi bylo předvedeno řešení minimalizace nákladů na provoz skladu a ukázáno, že i pro $(0,K_1,\ldots,K_n)$-konvenxí funkce platí téměř stejný výsledek za předpokladu, že jednotlivé náklady jsou na sobě nezávislé. U $(K_0,0,\ldots,0)$-konvexity jsem ukázali, že za předpokladu spojitosti a koercivity účelové funkce je řešení problému minimalizace nákladů na provoz skladu analogické řešení téže úlohy v jedné dimenzi, rozdílem je nahrazení podmínky $x < s$ podmínkou $x \in \sigma$. Zajímavější problém, tedy pro $\mathcal{K} = (K_0,K_1,\ldots,K_n)$ je v Gallego a Sethi (2005) uveden jako otevřený.

K-konvexita má podobnou definici jako konvexní funkce a zachovávají ji stejné operace, které zachovávají i konvexitu, ale nemá žádnou z pěkných vlastností konvexních funkcí, které jsou zmíněny v abstraktu.

Jednou z otázek, které se mohou nabízet, je, zda-li je potřeba zabývat se kvazikonvexními funkcemi, když byly objeveny funkce kvazi-invexní. Stejně jako je tomu mezi konvexními a kvazikonvexními funkcemi, kvazikonvexní funkce mají řadu vlastností, například konvexní úrovňové množiny, které kvazi-invexní funkce nemají. Na druhou stranu invexní funkce nabízejí možnost pro zkoumání zobecněných konvexních funkcí skrze funkci η.

Seznam použité literatury

Agrawal, A. a Boyd, S. (2020). Disciplined quasiconvex programming. *Optimization Letters*. doi: 10.1007/s11590-020-01561-8.

Barto, L. a Tůma, J. (2019). Lineární algebra. URL `http://www.karlin.mff.cuni.cz/~barto/LinAlg/skripta_la6.pdf`.

Bazaraa, M. S. (2006). *Nonlinear Programming: Theory and Algorithms*. Wiley, Hoboken, New Jersey. ISBN 978-0-471-48600-8.

Ben-Israel, A. a Mond, B. (1986). What is invexity? *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, **28**(1), 1–9. doi: 10.1017/S0334270000005142.

Boyd, S. (2004). *Convex Optimization*. Cambridge University Press, Cambridge, UK. ISBN 978-0-521-83378-3.

Colell, A. (1995). *Microeconomic Theory*. Oxford University Press, New York. ISBN 0-19-507340-1.

Dupačová, J. a Lachout, P. (2011). *Úvod do optimalizace*. Matfyzpress, Prague. ISBN 978-80-7378-176-7.

Eppstein, D. (2004). Quasiconvex programming. URL `http://library.msri.org/books/Book52/files/16epp.pdf`.

Gallego, G. a Sethi, S. (2005). K-Convexity in Rn. *Journal of Optimization Theory and Applications*, **127**, 71–88. doi: 10.1007/s10957-005-6393-4.

Hanson, M. A. (1981). On sufficiency of the Kuhn-Tucker conditions. *Journal of Mathematical Analysis and Applications*, **80**(2), 545–550. doi: 10.1016/0022-247x(81)90123-2.

Hanson, M. A. (1999). Invexity and the Kuhn–Tucker theorem. *Journal of Mathematical Analysis and Applications*, **236**(2), 594–604. doi: 10.1006/jmaa.1999.6484.

Lachout, P. (2011). Konvexita v konečné dimenzi. URL `http://www.karlin.mff.cuni.cz/~lachout/Vyuka/TEXTY/111016-Konvexita-skripta.pdf`.

Rockafellar, R. T. (1997). *Convex analysis*. Princeton University Press, Princeton, New Jersey. ISBN 0-691-08069-0.

Scarf, H. (2005). The Optimality of (S, s) Policies in the Dynamic Inventory Problem. In *Herbert Scarf's Contributions to Economics, Game Theory and Operations Research*, pages 1–7. Palgrave Macmillan UK. doi: 10.1057/9781137024381_1.