UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Dualita geometrického programovania a jej možné analógie

BAKALÁRSKA PRÁCA

2013 Martin HURBAN

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Dualita geometrického programovania a jej možné analógie

BAKALÁRSKA PRÁCA

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	1114, Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. RNDr. Milan Hamala, CSc

Bratislava 2013 **Martin Hurban**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Martin Hurban
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský

Názov: Dualita geometrického programovania a jej možné analógie

Cieľ: Podľa schémy duality geometrického programovania pokúsiť sa navrhnúť analogické dvojice duálnych úloh a numericky ilustrovať výhodnosť riešenia príslušných duálnych úloh v porovnaní s náročnosťou riešenia primárnych úloh.

Vedúci: doc. RNDr. Milan Hamala, CSc.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.

Dátum zadania: 25.10.2012

Dátum schválenia: 03.11.2012 doc. RNDr. Margaréta Halická, CSc.
garant študijného programu

....
študent vedúci práce

Poďakovanie Touto cestou sa chcem poďakovat svojmu vedúcemu bakalárskej práce Doc. RNDr. Milanovi Hamalovi, CSc. za ochotu, pomoc, odborné aj štylistické rady a podnetné pripomienky, ktoré mi veľmi pomohli pri, snáď zrozumiteľnom, spísaní tejto práce. Ďakujem aj Márii Mészárosovej za veľa čiarok, ktoré v budúcnosti opraví. Ďakujem aj svojej rodine a priateľom za ich trpezlivosť a podporu.

Abstrakt v štátnom jazyku

HURBAN, Martin: Dualita geometrického programovania a jej možné analógie [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. RNDr. Milan Hamala, CSc., Bratislava, 2013, 42 s.

V našej práci využívame teóriu duality konvexného programovania, aby sme k primárnej úlohe zostrojili duálnu úlohu, ktorá bude mať jednoduchšiu štruktúru ako primárna úloha. Pomocou teórie triviálnej duality pre úlohy bez ohraničení sa snažíme určiť primárne úlohy, ku ktorým sa dá zostrojiť duálna úloha s jednoduchšou štruktúrou. Navrhli sme "všeobecný model", ktorý spája všetky príklady uvedené v práci a tak zahŕňa niekoľko rôznych typov primárnych úloh. Okrem toho všeobecný model obsahuje aj parameter, ktorý umožňuje hľadanie nových primárnych úloh. Ďalej sme sa v práci snažili určiť všetky účelové funkcie, ktoré do "všeobecného modelu patria". Podarilo sa nám ukázať, že účelové funkcie spĺňajúce náš model generujú úlohy minimalizácie zovšeobecnených kladných polynómov -geometrické programovanie, p-noriem a mocninových priemerov. Ďalej sa nám podarilo určiť nutné podmienky pre úlohy, ktorým možno vytvoriť duálnu úlohu pomocou schémy v našom modeli. Tieto podmienky spĺňajú len už vymenované triedy úloh. Taktiež sme v troch vzorových príkladoch popísali postup, akým možno získať riešenia primárnej úlohy pomocou vyriešenia duálnej úlohy a vyriešenia sústavy lineárnych rovníc.

Kľúčové slová: triviálna dualita, geometrické programovanie, konvexné programovaie, Wolfeho dualita, Cauchyho multiplikatívna rovnica

Abstract

HURBAN, Martin: Index Policies for Dynamic and Stochastic Problems [Bachelor Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. RNDr. Milan Hamala, CSc., Bratislava, 2013, 42p.

In the duality theory of convex programming, there is always a dual problem corresponding primal one. In our thesis we are using the fact, that dual problem is sometimes simpler the primal. Using trivial duality theory for unconstrained problems we are trying to determine primal problems, for which simpler dual problems can be constructed. We designed "general model", in which all problems listed in this thesis are included and so this model contains different types of primal problems. Moreover general model contains functional parameter, which enables search for new primal problems. In work we tried to determine all objective functions satisfying general model. We showed, that objective functions satisfying our model generate problems of minimising posynomials - Geometric programming, p-norms and power means. We were also capable of determining necessary conditions for primal problems, to which dual problems by our scheme can be constructed. The necessary conditions are fullfilled only by mentioned classes of problems. Also we had described method of solving primal problem by solving dual one and system of linear equations in three model problems.

Keywords: Trivial duality, Geometric programming, Convex programming, Wolfe duality, Cauchy multiplicative equation

Obsah

Úvod **8**

1 Príklady riešenia úloh pomocou duality **9**

1.1 Príklad 1 - Geometrické programovanie 9

1.2 Príklad 2 12

1.3 Príklad 3 16

2 Všeobecný model **20**

2.1 Rozbor príkladov z kapitoly 1 22

2.2 Zhrnutie modelu 23

3 Postačujúce podmienky funkcií vo všeob. modeli **24**

3.1 Logaritmická funkcia 24

3.2 Mocninová funkcia 25

4 Nutné podmienky funkcií vo všeob. modeli **29**

4.1 Vlastnosti funkcií v modeli 29

4.2 Odvodenie nutných podmienok I.časť 30

4.3 Cauchyho multiplikatívna rovnica v prirodzených číslach 32

4.4 Odvodenie nutných podmienok II.časť 36

Záver **39**

Zoznam použitej literatúry **40**

Príloha A (Dôkaz tvrdenia zo str. 17 a 27) **42**

Úvod

V teórii konvexného programovania ku každej úlohe na viazaný extrém, nazývanej primárna úloha, je definovaná tzv. duálna úloha. Premennými v duálnej úlohe sú tzv. Lagrangeove premenné, z ktorých každá prislúcha niektorému ohraninčeniu duálnej úlohy.

Vo svojej monografii [3] z r. 1967 Duffin-Peterson-Zener skúmali špeciálnu dvojicu duálnych úloh, u ktorej bolo možné skonštruovať duálnu úlohu aj k primárnej úlohe na voľný extrém, t.j. úlohu bez ohraničení. V takejto úlohe nevystupujú Lagrangeove premenné. Príslušnú teóriu nazvali "Geometrické programovanie".

Vo svojej monografii [6] z r. 1972 Hamala ukázal, že Geometrické programovanie je špeciálnym prípadom Wolfeho duality, čo viedlo k definovaniu Wolfeho duálnej úlohy pre primárnu úlohu bez ohraničení. Príslušnú teóriu nazval "triviálna dualita". Na základe triviálnej duality Hamala uviedol ďalšie dve dvojice duálnych úloh analogických geometrickému programovaniu.

Uvedené skutočnosti viedli k prirodzenej otázke, či existujú ďalšie zmysluplné dvojice duálnych úloh odvodené z triviálnej duality. **Cieľom** predloženej bakalárskej práce bolo dať odpoveď na uvedenú otázku. Pri skúmaní príkladov z [6] sa nám podarilo uvedenú otázku nielen zodpovedať, ale sformulovali sme a dokázali nutné a postačujúce podmienky pre konštrukciu uvedených dvojích úloh.

Naša práca práca pozostáva zo štyroch kapitol, ktoré nám postupne dajú návod ako zostrojiť triviáne duálnu úlohu, ukážu nám postup ako využiť dualitu pri riešení optimalizačných úloh a na záver presne vymedzia úlohy, ku ktorým sa naším prístupom dá vytvoriť duálna úloha s jednoduchšou štruktúrou.

V 1.kapitole uvedime príklady z [6]. Na základe ich rozboru navrhneme v 2. kapitole všeobecný model, ktorý zahŕňa všetky tri príklady z kapitoly 1. Nosnou časťou práce sú kapitoly 3 a 4, kde uvedieme nutné a postačujúce podmienky pre konštrukciu dvojice duálnych úloh.

Motiváciou k výberu tejto témy boli prednášky z nelineárneho programovania [5], túžba po hlbšom chápaní duality v nelineárnom programovaní a osvojenie si vedomostí zo spomínaných prednášok.

1 Príklady riešenia úloh pomocou duality

V tejto kapitole uvedieme príklady transformácie jednotlivých úloh s použitím triviálnej duality a postup ich riešenia. Primárnou úlohou bude úloha na voľný extrém, kde f je konvexná, $f : \mathbb{R}^m \to \mathbb{R}$ a $f \in C^1$:

$$Min \{f(z) \mid z \in \mathbb{R}^m\}. \tag{1}$$

Pomocou teórie triviálnej duality - duality bez ohraničení zostrojíme Wolfeho duálnu úlohu v tvare:

$$Max \{f(y) \mid \nabla f(y) = 0_m\} \tag{2}$$

Príklady v tejto kapitole pochádzajú z knihy [6, str. 108 - 116]. Príkladom geometrického programovania aj s ohraničeniami sa podrobne zaoberá celá kniha [3].

1.1 Príklad 1 - Geometrické programovanie

Geometrické programovanie v prípade bez ohraničení rieši úlohu minimalizovať tzv. zovšeobecnený kladný polynóm teda $g(t) = \sum_{i=1}^{n} c_i \prod_{j=1}^{m} t_j^{a_{ij}}$, kde $c_i > 0$ a $a_{ij} \in \mathbb{R}$ s definičným oborom $\mathbb{R}_+^m$, teda $\forall j = 1, 2 \ldots m : t_j > 0$. Úloha, ktorú chceme riešiť, je vlastne úlohou na voľný extrém a má tvar :

$$Min \left\{ g(z) = \sum_{i=1}^{n} c_i \prod_{j=1}^{m} t_j^{a_{ij}} \mid t \in \mathbb{R}_+^m \right\}, \text{ kde } c_i > 0 \text{ a } a_{ij} \in \mathbb{R}. \tag{3}$$

Keďže daný zovšeobecnený polynóm vo všeobecnosti nie je konvexná funkcia, zvolíme si substitúciu $t_j = e^{z_j} \Leftrightarrow z_j = ln(t_j)$. Po tejto substitúcii je g v premenných z_j konvexná. To možno ľahko vidieť z toho, že $e^{\sum_{j=1}^{m} a_{ij} z_j}$ je zloženie rastúcej konvexnej funkcie a konvexnej funkcie, čo je konvexná funkcia a súčet konvexých funkcií je tiež konvexná funkcia. Definičným oborom g je celé $\mathbb{R}^m$.

$$Min \left\{ g(t) = \sum_{i=1}^{n} c_i e^{\sum_{j=1}^{m} a_{ij} z_j} \mid z \in \mathbb{R}^m \right\}, \text{ kde } c_i > 0 \text{ a } a_{ij} \in \mathbb{R}. \tag{4}$$

Ďalej označíme $x_i := \sum_{j=1}^{m} a_{ij} z_j$ pre $\forall i = 1 \ldots n$. Ku funkcii g pridáme ešte vonkajšiu transformáciu a tak vytvoríme funkciu $f(z) := ln(g(z))$. Dôkaz, že táto účelová funkcia je konvexná, je uvedený v [6, str. 110]. Takto upravená úloha má tvar:

$$Min\left\{f(z)=ln\left(\sum_{i=1}^{n}c_ie^{x_i}\right)\mid x_i=\sum_{j=1}^{m}a_{ij}z_j,z\in\mathbb{R}^m\right\}, \text{ kde } c_i>0 \text{ a } a_{ij}\in\mathbb{R}. \quad (5)$$

K tejto úlohe vytvoríme Wolfeho duálnu úlohu, ktorú možno zostrojiť, pretože účelová funkcia je konvexná. Prechod je analogický ako z (1) na (2).

Počítajme parciálnu deriváciu podľa z_j:

$$\frac{\partial f}{\partial z_j}=\frac{1}{g(z)}\left[\sum_{i=1}^{n}c_ie^{x_i}\cdot a_{ij}\right]=0 \text{ pre } \forall j=1\dots m. \quad (6)$$

Zavedieme substitúciu

$$\delta_i:=\frac{c_ie^{x_i}}{g(z)}>0,\ \forall i=1\dots n. \quad (7)$$

Táto substitúcia zužuje definičný obor duálnej úlohy na $\delta\in\mathbb{R}^m_+$. Takto zúžený definičný obor ostáva stále otvorenou množinou a nemení charakter riešenej úlohy. Po tejto substitúcii (6) nadobúda jednoduchší tvar :

$$\sum_{i=1}^{n}\delta_ia_{ij}=0, \quad (8)$$

čo vo vektorovom tvare možno zapísať ako $A\delta=0_m$, kde $A=(a_{ij})$.
Zo substitúcie (7) vyplýva vzťah:

$$\sum_{i=1}^{n}\delta_i=1, \quad (9)$$

čo vo vektorovom tvare možno zapísať ako $e^T\delta=1$.
Teraz sa nám podarilo upraviť ohraničenia úlohy duálnej k (5) na vektorový tvar :

$$\{A\delta=0_m,e^T\delta=1,\delta>0_m\}. \quad (10)$$

Ostáva nám len nahradiť v účelovej funkcii $f(z)=ln\left(\sum_{i=1}^{n}c_ie^{x_i}\right)$ premennnú z novou prememnnou δ. Vychádzajme zo substitúcie (7), v tvare:

$$\delta_ig(z)=c_ie^{x_i},\ \forall i=1\dots n. \quad (11)$$

Pravá aj ľavá strana sú kladné. Preto možno umocnením na δ_i a následným zlogaritmovaním vzťahov (11) získať :

$$\delta_i ln\left(\delta_i\right) + \delta_i f\left(z\right) = \delta_i ln\left(c_i\right) + \delta_i\ x_i,\ \forall i = 1 \ldots n. \tag{12}$$

Teraz možno všetky rovnice (12) sčítať. Pre koeficient pri člene $f\left(z\right)$ možno využiť vzťah (9), čo po úpravách dáva:

$$f\left(z\right) = \sum_{i=1}^{n} \delta_i ln\left(\frac{c_i}{\delta_i}\right) + \sum_{i=1}^{n} \delta_i x_i. \tag{13}$$

Ukážeme, že posledný člen je nulový. Rozviňme zápis pre x_i. Dostávame $\sum_{i=1}^{n} \delta_i x_i = \sum_{i=1}^{n} \delta_i \left(\sum_{j=1}^{m} a_{ij} z_j\right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \delta_i a_{ij}\right)$ a vďaka (8) je celý výraz rovný 0. A vzťah (13) možno napísať ako :

$$f\left(z\right) = \sum_{i=1}^{n} \delta_i ln\left(\frac{c_i}{\delta_i}\right). \tag{14}$$

Spojením vzťahov (8), (9), (14) a označením $v\left(\delta\right) := \prod_{i=1}^{n} \left(\frac{c_i}{\delta_i}\right)^{\delta_i}$ získavame duálnu úlohu k úlohe (5):

$$Max \left\{ ln\left(v\left(\delta\right)\right) = \sum_{i=1}^{n} \delta_i ln\left(\frac{c_i}{\delta_i}\right) \mid \sum_{i=1}^{n} \delta_i a_{ij} = 0,\text{ pre } \forall j = 1 \ldots m.\ , \sum_{i=1}^{n} \delta_i = 1, \delta \in \mathbb{R}_+^m \right\}. \tag{15}$$

Prípadne duálnu úlohu k úlohe (4) :

$$Max \left\{ v\left(\delta\right) = \prod_{i=1}^{n} \left(\frac{c_i}{\delta_i}\right)^{\delta_i} \mid \sum_{i=1}^{n} \delta_i a_{ij} = 0,\text{ pre } \forall j = 1 \ldots m.\ , \sum_{i=1}^{n} \delta_i = 1, \delta \in \mathbb{R}_+^m \right\}. \tag{16}$$

Získaná duálna úloha (15) má jednoduchšiu štruktúru ako primárna úloha (5) a to najmä vďaka:

- Lineárnym ohraničeniam a
- Separovateľnej a konkávnej účelovej funkcii.

Zhrnutie:

1. Chceme riešiť úlohu (3).
2. Namiesto nej vyriešime jednoduchšiu úlohu (15). Získame $\hat{\delta} \in \mathbb{R}^m_+$ a $ln\left(v\left(\hat{\delta}\right)\right)$.
3. Aplikovaním teórie triviálnej duality priamo získavame hodnotu optima úlohy (3): $f\left(\hat{z}\right) = ln\left(v\left(\hat{\delta}\right)\right)$.
4. Použitím substitučných vzťahov (7) dopočítam hodnotu argumentu optima, teda $\hat{z} \in \mathbb{R}^m$.
5. Posledným krokom je návrat k pôvodným premenným zo vzťahu $\hat{t}_j = e^{\hat{z}_j}$.

Nasleduje úprava vzťahov (7), ktorá nás dovedie k úlohe vyriešiť sústavu n lineárnych rovníc o m neznámych $\hat{z}_j$:

$$\begin{aligned} \hat{\delta}_i g\left(\hat{z}\right) &= c_i e^{\hat{x}_i}, \text{ pre } \forall i = 1 \dots n \\ ln\left(\hat{\delta}_i\right) + f\left(\hat{z}\right) &= ln\left(c_i\right) + \sum_{j=1}^{m} a_{ij}\hat{z}_j, \\ \text{keďže } f\left(\hat{z}\right) = ln\left(v\left(\hat{\delta}\right)\right) &\text{ možno napísať:} \\ \sum_{j=1}^{m} a_{ij}\hat{z}_j = ln\left(v\left(\hat{\delta}\right)\right) + ln\left(\frac{\hat{\delta}_i}{c_i}\right)&, \text{ pre } \forall i = 1 \dots n. \end{aligned} \tag{17}$$

Posledný vzťah jasne určuje sústavu n lineárnych rovníc o m neznámych $\hat{z}_j$. Každé riešenie danej sústavy možno spätne transformovať do pôvodnej premennej t, a to vzťahom $\hat{t}_j = e^{\hat{z}_j}$. To platí pre $\forall j = 1 \dots m$, potom každý vektor $\hat{t}$ je je riešením pôvodnej úlohy (3).

1.2 Príklad 2

Tento príklad bude ďalšou ukážkou ako sa dá pomocou transformácie a dualizácie úlohy prísť k jej riešeniu. Úloha, ktorú budeme riešiť, je v [6, str. 108] riešená pre druhú mocninu. My budeme riešiť úlohu so štvrtou mocninou. Chceme zistiť riešenie nekonvexnej úlohy :

$$Min\left\{ g\left(t\right)=\sum_{i=1}^{n}\left(c_i+\sum_{j=1}^{m}a_{ij}t_j^{\frac{1}{4}}\right)^4 \mid t\in\mathbb{R}_+^m\cup\{0\}\right\}. \tag{18}$$

Danú úlohu upravíme na konvexnú pomocou substitúcie $t_j := z_j^4$. Takto upravená úloha je konvexná v premennej z :

$$Min\left\{ g\left(z\right)=\sum_{i=1}^{n}\left(c_i+\sum_{j=1}^{m}a_{ij}z_j\right)^4 \mid z\in\mathbb{R}_+^m\cup\{0\}\right\}. \tag{19}$$

Zaveďme označenie analogické k predchádzajúcemu príkladu, a to $x_i := \sum_{j=1}^{m} a_{ij}z_j$. Ďalej pridajme vonkajšiu inverznú transformáciu. Takto získaná úloha má tvar :

$$Min\left\{ f\left(z\right)=\left[\sum_{i=1}^{n}\left(c_i+x_i\right)^4\right]^{\frac{1}{4}} \mid x_i=\sum_{j=1}^{m}a_{ij}z_j, z\in\mathbb{R}_+^m\cup\{0\}\right\}. \tag{20}$$

Definičným oborom úlohy (20) budeme chápať množinu kde platí: $c_i + x_i > 0$ [1] . Dôvodom je zachovanie jednoznačnosti riešenia a vylúčenie triviálneho globálneho miminma $x_i + c_i = 0$. Ďalej platí, že takto tranformovaná účelová funkcia ostáva konvexná. Vyplýva to z toho, že pre $c_i + x_i > 0$ sa funkcia $f(z)$ správa ako p-norma, kde $p = 4$. A keďže je všeobecne známe, že každá norma je konvexná funkcia, aj naša účelová funkcia je teda konvexná.

Vďaka konvexnosti teraz môžme postupne zostrojiť k úlohe (20) duálnu úlohu. Rátajme parciálne derivácie podľa premenných z_j a položme ich rovné 0 :

$$\frac{\partial f\left(z\right)}{\partial z_j}=\frac{1}{4}\left[g\left(z\right)\right]^{\frac{-3}{4}}4\sum_{i=1}^{n}\left(c_i+x_i\right)^3a_{ij}=0,\text{ pre }\forall j=1\ldots m. \tag{21}$$

Zvolíme substitúciu, ktorá nás prenesie od premenných z k duálnym premenným δ:

$$\delta_i := \frac{\left(c_i+x_i\right)^3}{\left[\sum_{i=1}^{n}\left(c_i+x_i\right)^4\right]^{\frac{3}{4}}}\text{ , pre }\forall i=1\ldots n. \tag{22}$$

Táto substitúcia zužuje definičný obor duálnej úlohy na $\delta \in \mathbb{R}_+^m$. Takto zúžený definičný obor ostáva stále otvorenou množinou a nemení charakter riešenej úlohy. Po tejto substitúcii získal vzťah (21) tvar :

[1] Táto množina je otvorená, preto sa charakter úlohy nemení a ide o hľadanie voľného extrému.

$$\sum_{i=1}^{n} \delta_i a_{ij} = 0, \tag{23}$$

čo vo vektorovom tvare možno zapísať ako $A\delta = 0_m$, kde $A = (a_{ij})$. Zo substitúcie (22), vyplýva vzťah:

$$\sum_{i=1}^{n} \delta_i^{\frac{4}{3}} = 1, \tag{24}$$

čo vo vektorovom tvare možno zapísať ako $e^T \delta^{\frac{4}{3}} = 1$ Teraz sa nám podarilo upraviť ohraničenia úlohy duálnej k (20) na vektorový tvar:

$$\{A\delta = 0_m, e^T \delta^{\frac{4}{3}} = 1, \delta > 0_m\}. \tag{25}$$

Ostáva nám len nahradiť v účelovej funkcii $f(z) = \left[\sum_{i=1}^{n} (c_i + x_i)^4\right]^{\frac{1}{4}}$ prememnnú z novou prememnnou δ. Vychádzajme zo substitúcie (22), v tvare:

$$\delta_i \left(g(z)^{\frac{3}{4}} \right) = (c_i + x_i)^3 . \tag{26}$$

Prenásobením každej rovnice postupne $(c_i + x_i)$ a ich sčítaním možno získať:

$$(g(z))^{\frac{3}{4}} \sum_{i=1}^{n} \delta_i (c_i + x_i) = g(z) . \tag{27}$$

Prenásobme rovnicu $(g(z))^{-\frac{3}{4}}$ a na pravej strane získame vyjadrenie pre $f(z)$:

$$\sum_{i=1}^{n} \delta_i (c_i + x_i) = f(z) \tag{28}$$

Ukážeme, že z výrazu možno vylúčiť x_i. Rozviňme zápis pre $\delta_i x_i$. Dostávame:

$$\sum_{i=1}^{n} \delta_i x_i = \sum_{i=1}^{n} \delta_i \left(\sum_{j=1}^{m} a_{ij} z_j \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \delta_i a_{ij} \right) \tag{29}$$

a vďaka (8) je celý výraz rovný 0. Vzťah (28) možno napísať ako :

$$f(z) = \sum_{i=1}^{n} c_i \delta_i . \tag{30}$$

Spojením vzťahov (23), (24), (30) a označením $v(\delta) := \left(\sum_{i=1}^{n} c_i \delta_i\right)^4$ získavame duálnu úlohu k úlohe (20):

$$Max\left\{(v(\delta))^{\frac{1}{4}}=\sum_{i=1}^{n}c_i\delta_i \mid \sum_{i=1}^{n}\delta_i a_{ij}=0,\text{ pre }\forall j=1\dots m.\ ,\sum_{i=1}^{n}\delta_i^{\frac{4}{3}}=1,\delta\in\mathbb{R}_+^m\right\}. \tag{31}$$

Prípadne duálnu úlohu k úlohe (19) :

$$Max\left\{v(\delta)=\left(\sum_{i=1}^{n}c_i\delta_i\right)^4 \mid \sum_{i=1}^{n}\delta_i a_{ij}=0,\text{ pre }\forall j=1\dots m.\ ,\sum_{i=1}^{n}\delta_i^{\frac{4}{3}}=1,\delta\in\mathbb{R}_+^m\right\}. \tag{32}$$

Získaná duálna úloha (31) má jednoduchšiu štruktúru ako primárna úloha (20), a to najmä vďaka:

- ohraničeniam, ktoré sú okrem jedného lineárne
- lineárnej účelovej funkcii.

Zhrnutie:

1. Chceme riešiť úlohu (18).
2. Namiesto nej vyriešime jednoduchšiu úlohu (31). Získame $\hat{\delta}\in\mathbb{R}_+^m$ a $\left(v\left(\hat{\delta}\right)\right)^{\frac{1}{4}}$.
3. Aplikovaním teórie triviálnej duality priamo získavame hodnotu optima úlohy (18): $f(\hat{z})=\left(v\left(\hat{\delta}\right)\right)^{\frac{1}{4}}$.
4. Použitím substitučných vzťahov (22) dopočítame hodnotu argumentu optima teda $\hat{z}\in\mathbb{R}^m$.
5. Posledným krokom je návrat k pôvodným premenným zo vzťahu $\hat{t}_j=\hat{z}_j{}^4$.

Nasleduje úprava vzťahov (22), ktorá nás dovedie k úlohe vyriešiť sústavu n lineárnych rovníc o m neznámych $\hat{z}_j$:

$$\begin{aligned}(c_i+\hat{x}_i)^3&=\hat{\delta}_i\left(g(\hat{z})\right)^{\frac{3}{4}},\\ (c_i+\hat{x}_i)&=\hat{\delta}_i^{\frac{1}{3}}f(\hat{z}),\\ \text{keďže } f(\hat{z})=\left(v\left(\hat{\delta}\right)\right)^{\frac{1}{4}}&=c^T\hat{\delta}\text{ možno napísať:}\\ \sum_{j=1}^{m}a_{ij}\hat{z}_j&=\hat{\delta}_i^{\frac{1}{3}}c^T\hat{\delta}-c_i,\text{ pre }\forall i=1\dots n.\end{aligned} \tag{33}$$

Posledný vzťah jasne určuje sústavu n lineárnych rovníc o m neznámych $\hat{z}_j$. Každé riešenie danej sústavy možno spätne transformovať do pôvodnej premennej t, a to vzťahom $\hat{t_j} = \hat{z_j}^4$. To platí pre $\forall j = 1 \ldots m$, potom každý vektor $\hat{t}$ je je riešením pôvodnej úlohy (18).

1.3 Príklad 3

Tento príklad bude poslednou ukážkou toho ako sa dá pomocou transformácie a dualizácie úlohy prísť k jej riešeniu. Úloha, ktorú budeme riešiť, je v [6, str. 116] uvedená ako cvičenie 7. My si príklad podrobne rozoberieme. Chceme zistiť riešenie nekonvexnej úlohy:

$$Min \left\{ g\left(t\right) = \sum_{i=1}^{n} \left(c_i + \sum_{j=1}^{m} a_{ij} t_j^{-1} \right)^{-1} \mid t \in \mathbb{R}_+^m \right\}. \tag{34}$$

Danú úlohu upravíme na konvexnú pomocou substitúcie $t_j := z_j^{-1}$. Takto upravená úloha je konvexná v premennej z :

$$Min \left\{ g\left(z\right) = \sum_{i=1}^{n} \left(c_i + \sum_{j=1}^{m} a_{ij} z_j \right)^{-1} \mid z \in \mathbb{R}_+^m \right\}. \tag{35}$$

Zaveďme označenie analogické k predchádzajúcemu príkladu, a to $x_i := \sum_{j=1}^{m} a_{ij} z_j$. Ďalej pridajme vonkajšiu inverznú transformáciu. Keďže hyperbola je klesajúca funkcia, mení sa smer extremalizácie na opačný. Získaná úloha má tvar :

$$Max \left\{ f\left(z\right) = \left[\sum_{i=1}^{n} \left(c_i + x_i\right)^{-1} \right]^{-1} \mid x_i = \sum_{j=1}^{m} a_{ij} z_j, z \in \mathbb{R}_+^m \right\}. \tag{36}$$

Definičným oborom budeme chápať množinu kde platí: $c_i + x_i > 0$ [2] . Dôvodom je zachovanie jednoznačnosti riešenia. Taktiež sa musíme zúžiť na oblasť, kde je hyperbola monotónna. Ďalej platí, že zložená funkcia $f\left(z\right) = \left[\sum_{i=1}^{n} \left(c_i + x_i\right)^{-1}\right]^{-1}$ je konkávna (dôkaz je v Prílohe A).

Keďže úloha je maximalizačná s konkávnou účelovou funkciou možno postupne zostrojiť k úlohe (36) Wolfeho duálnu úlohu. Rátajme parciálne derivácie podľa premenných z_j a položme ich rovné 0 :

[2] Táto množina je otvorená, preto stále ide o hľadanie voľného extrému.

$$\frac{\partial f(z)}{\partial z_j} = -\left[g(z)\right]^{-2} \sum_{i=1}^{n} (c_i + x_i)^{-2} a_{ij} = 0, \text{ pre } \forall j = 1 \dots m. \tag{37}$$

Zvolíme substitúciu, ktorá nás prenesie od premenných z k duálnym premenným δ:

$$\delta_i := \left[g(z)\right]^{-2} \cdot (c_i + x_i)^{-2} \quad \text{, pre } \forall i = 1 \dots n. \tag{38}$$

Táto substitúcia zužuje definičný obor duálnej úlohy na $\delta \in \mathbb{R}_+^m$. Takto zúžený definičný obor ostáva stále otvorenou množinou a nemení charakter riešenej úlohy. Po tejto substitúcii získal vzťah (37) tvar:

$$\sum_{i=1}^{n} \delta_i a_{ij} = 0, \tag{39}$$

čo vo vektorovom tvare možno zapísať ako $A\delta = 0_m$, kde $A = (a_{ij})$. Zo substitúcie (38) vyplýva vzťah:

$$\sum_{i=1}^{n} \delta_i^{\frac{1}{2}} = 1, \tag{40}$$

čo vo vektorovom tvare možno zapísať ako $e^T \delta^{\frac{1}{2}} = 1$. Teraz sa nám podarilo upraviť ohraničenia úlohy duálnej k (36) na vektorový tvar:

$$\{A\delta = 0_m, e^T \delta^{\frac{1}{2}} = 1, \delta > 0_m\}. \tag{41}$$

Ostáva nám len nahradiť v účelovej funkcii $f(z) = \left[\sum_{i=1}^{n} (c_i + x_i)^{-1}\right]^{-1}$ premennú z novou prememnnou δ. Vychádzajme zo substitúcie (38) v tvare:

$$\delta_i^{\frac{1}{2}} = \left(g(z) \cdot (c_i + x_i)\right)^{-1}. \tag{42}$$

Postupným prenásobením každej rovnice výrazom $(c_i + x_i)\, \delta_i^{\frac{1}{2}}$ možno získať:

$$(c_i + x_i)\, \delta_i = f(z)\, \delta_i^{\frac{1}{2}}. \tag{43}$$

Sčítajme všetky rovnice a vďaka (40) na pravej strane získame vyjadrenie pre $f(z)$:

$$\sum_{i=1}^{n} \delta_i (c_i + x_i) = f(z). \tag{44}$$

Ukážeme, že z výrazu možno vylúčiť x_i. Rozviňme zápis pre $\delta_i x_i$. Dostávame:

$$\sum_{i=1}^{n} \delta_i x_i = \sum_{i=1}^{n} \delta_i \left(\sum_{j=1}^{m} a_{ij} z_j \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \delta_i a_{ij} \right) \tag{45}$$

a vďaka (39) je celý výraz rovný 0. Vzťah (44) možno napísať ako:

$$f(z) = \sum_{i=1}^{n} c_i \delta_i. \tag{46}$$

Spojením vzťahov (39), (40), (46) a označením $v(\delta) := \left(\sum_{i=1}^{n} c_i \delta_i\right)^4$ získavame duálnu úlohu k úlohe (36):

$$Min \left\{ (v(\delta))^{-1} = \sum_{i=1}^{n} c_i \delta_i \mid \sum_{i=1}^{n} \delta_i a_{ij} = 0, \text{ pre } \forall j = 1 \ldots m. , \sum_{i=1}^{n} \delta_i^{\frac{1}{2}} = 1, \delta \in \mathbb{R}_+^m \right\}, \tag{47}$$

prípadne duálnu úlohu k úlohe (35):

$$Max \left\{ v(\delta) = \left(\sum_{i=1}^{n} c_i \delta_i \right)^{-1} \mid \sum_{i=1}^{n} \delta_i a_{ij} = 0, \text{ pre } \forall j = 1 \ldots m. , \sum_{i=1}^{n} \delta_i^{\frac{1}{2}} = 1, \delta \in \mathbb{R}_+^m \right\}. \tag{48}$$

Získaná duálna úloha (47) má jednoduchšiu štruktúru ako primárna úloha (36) a to najmä vďaka:

- ohraničeniam, ktoré sú okrem jedného lineárne,
- lineárnej účelovej funkcii.

Zhrnutie:

1. Chceme riešiť úlohu (34).
2. Namiesto nej vyriešime jednoduchšiu úlohu (35). Získame $\hat{\delta} \in \mathbb{R}_+^m$ a $\left(v\left(\hat{\delta}\right)\right)^{-1}$.
3. Aplikovaním teórie triviálnej duality priamo získavame hodnotu optima úlohy (34): $f(\hat{z}) = \left(v\left(\hat{\delta}\right)\right)^{-1}$.
4. Použitím substitučných vzťahov (38) dopočítame hodnotu argumentu optima teda $\hat{z} \in \mathbb{R}^m$.

5. Posledným krokom je návrat k pôvodným premenným zo vzťahu $\hat{t}_j = \hat{z}_j{}^{-1}$.

Nasleduje úprava vzťahov (38), ktorá nás dovedie k úlohe vyriešiť sústavu n lineárnych rovníc o m neznámych $\hat{z}_j$.

$$\begin{aligned} \left(c_i + \hat{x}_i\right)^2 \hat{\delta}_i &= \left(g\left(\hat{z}\right)\right)^{-2}, \\ \left(c_i + \hat{x}_i\right) \hat{\delta}_i^{\frac{1}{2}} &= f\left(\hat{z}\right). \\ \text{Keďže } f\left(\hat{z}\right) = \left(v\left(\hat{\delta}\right)\right)^{-1} &= c^T\hat{\delta} \text{ možno napísať:} \\ \hat{x}_i &= \hat{\delta}_i^{-\frac{1}{2}} c^T \hat{\delta} - c_i \\ \sum_{j=1}^{m} a_{ij}\hat{z}_j &= \hat{\delta}_i^{-\frac{1}{2}} c^T \hat{\delta} - c_i, \text{ pre } \forall i = 1 \ldots n. \end{aligned} \tag{49}$$

Posledný vzťah jasne určuje sústavu n lineárnych rovníc o m neznámych $\hat{z}_j$. Každé riešenie danej sústavy možno spätne transformovať do pôvodnej premennej t, a to vzťahom $\hat{t_j} = \hat{z}_j{}^{-1}$. To platí pre $\forall j = 1 \ldots m$. Potom každý vektor t je je riešením pôvodnej úlohy (34).

2 Všeobecný model

Na základe príkladov v predchádzajúcej kapitole sa možno pokúsiť o ich zovšeobecnenie a zhrnutie všetkých príkladov pod jeden model. Uvažujme 2 množiny:

$$\begin{aligned} T &= \{t \in \mathbb{R} \mid t_- < t < t_+\}, \\ Z &= \{z \in \mathbb{R} \mid z_- < z < z_+\}, \end{aligned} \tag{50}$$

kde t_-, z_- môžu nadobúdať aj nevlastnú hodnotu $(-\infty)$ a t_+, z_+ môžu nadobúdať nevlastnú hodnotu $(+\infty)$. Preto T, Z sú buď konečné intervaly, polpriamky alebo celé priamky. Zvoľme funkcie:

$$\begin{aligned} \varphi : T \to Z \ , \ \varphi \in C^1, \\ \psi : Z \to T \ , \ \psi \in C^1 \end{aligned} \tag{51}$$

tak, aby platilo, že sú rýdzomonotónne a vzájomne inverzné, t.j.

$$\begin{aligned} \forall z \in Z \subset \mathbb{R} : \varphi\left(\psi\left(z\right)\right) = z \\ \forall t \in T \subset \mathbb{R} : \psi(\varphi(t)) = t. \end{aligned} \tag{52}$$

Predpokladajme ďalej, že ψ je konvexná[3]. Potom platí, že φ ako jej inverz je konkávna.

Chceme riešiť úlohu :

$$Min \left\{ g\left(t\right) = \sum_{i=1}^{n} \psi \left[c_i + \sum_{j=1}^{m} a_{ij} \varphi\left(t_j\right) \right] \middle| \ t \in T^m \right\}, \text{ kde } c_i, \ a_{ij} \in \mathbb{R}. \tag{53}$$

Zavedieme substitúciu $t_j = \psi\left(z_j\right) \Leftrightarrow z_j = \varphi\left(t_j\right)$. Po tejto substitúcii zapíšeme úlohu (53) v konvexnom tvare:

$$Min \left\{ g\left(z\right) = \sum_{i=1}^{n} \psi \left[c_i + \sum_{j=1}^{m} a_{ij} z_j \right] \mid z \in Z^m \right\}, \text{ kde } c_i, \ a_{ij} \in \mathbb{R}. \tag{54}$$

Pridáme vonkajšiu transformáciu[4] , ktorá ak je φ rastúca dáva:

[3] Tento predpoklad nie je v ďalšom odvodení nevyhnutný a možno sa bez neho zaobísť, avšak ako sa neskôr ukáže náš model budú spĺňať len konvexné funkcie ψ.

[4] Treba povedať, že aby bolo možné zaviesť vonkajšiu trasformáciu, musí obraz funkcie $g\left(z\right)$ ležať v definičom obore φ. Z toho dôvodu musí byť množina T tvaru $\langle t_+; \infty \rangle$, $\langle -\infty; t_- \rangle$ alebo ich zjednotením, pričom $t_+ \geq 0$ a $t_- \leq 0$.

$$Min\left\{f\left(z\right)=\varphi\left(g\left(z\right)\right)\mid z\in Z^{m},\right\} \tag{55}$$

alebo ak je φ klesajúca dáva:

$$Max\left\{f\left(z\right)=\varphi\left(g\left(z\right)\right)\mid z\in Z^{m}\right\}. \tag{56}$$

Ak sú úlohy (55) a (56) úlohami konvexného programovania, t.j. ak je účelová funkcia v (55) konvexná, alebo ak je účelová funkcia v (56) konkávna, tak možno k obom zostrojiť Wolfeho duálne úlohy: pre úlohu (55) v tvare:

$$Max\left\{f\left(y\right)\mid\nabla f\left(y\right)=0_{m}\right\} \tag{57}$$

a pre úlohu (56) v tvare:

$$Min\left\{f\left(y\right)\mid\nabla f\left(y\right)=0_{m}\right\}. \tag{58}$$

Počítajme parciálnu deriváciu podľa z_j:

$$\frac{\partial f}{\partial z_j}=\varphi'\left(g\left(z\right)\right)\cdot\frac{\partial g}{\partial z_j}=\varphi'\left(g\left(z\right)\right)\cdot\sum_{i=1}^{n}\psi'\left(c_i+x_i\right)a_{ij}=0\ ,j=1,2,\ldots m, \tag{59}$$

kde $x_i=\sum_{j=1}^{m}a_{ij}z_j$.
Zaveďme substitúciu :

$$\delta_i:=\varphi'\left(g\left(z\right)\right)\cdot\psi'\left(c_i+x_i\right)\ ,i=1,2,\ldots n. \tag{60}$$

Pre premenné δ_i platí $\delta_i>0$, pretože δ_i je súčinom derivácií dvoch rastúcich alebo dvoch klesajúcich funkcií. Okrem toho po tejto subsititúcii (59) nadobúda tvar:

$$\sum_{i=1}^{n}\delta_i a_{ij}=0\ ,\ j=1,2,\ldots m. \tag{61}$$

Ďalej potrebujeme nájsť funkciu $\omega:\mathbb{R}_+\rightarrow\mathbb{R}_+$, ktorá bude spĺňať vzťah:

$$\sum_{i=1}^{n}\omega\left[\delta_i\right]=1. \tag{62}$$

Táto funkcia je pre príklady 1, 2, 3 z predchádzajúcej kapitoply postupne: x, $x^{\frac{4}{3}}$, $x^{\frac{1}{2}}$.
V uvedených prípadoch ω zabraňuje neohraničenosti daných problémov.
Ak uspejeme v hľadaní ω, tak zo vzťahu (60) získame vyjadrenie :

$$f(z) = \Omega(\delta) . \tag{63}$$

V tomto vyjadrení Ω reprezentuje účelovú funkciu vo vzťahu (57).
Ak sa nám podarí nájsť funkcie ω, Ω a zaistiť konvexnosť úlohy (55), resp. konkávnosť úlohy (56), potom zostrojíme duálnu úlohu v tvare:

$$Max \left\{ \Omega(\delta) \mid \sum_{i=1}^{n} \delta_i a_{ij} = 0, \sum_{i=1}^{n} \omega[\delta_i] = 1 \ , j = 1, 2, \dots m, \right\}, \text{ kde } \delta > 0. \tag{64}$$

2.1 Rozbor príkladov z kapitoly 1

Pozrime sa, ako sú jedotlivé príklady z predchádzajúcej kapitoly obsiahnuté naším všeobecným modelom.

Príklad 1.: Geometrické programovanie

$$T = \mathbb{R}^+ \qquad Z = \mathbb{R}$$
$$\varphi(x) = ln(x) \qquad \psi(x) = e^x$$
$$\omega(x) = x$$
$$\Omega(\delta) = \textstyle\sum_{i=1}^{n} \delta_i \cdot ln\left(\frac{c_i}{\delta_i}\right)$$

Príklad 2.: modifikácia príkladu 2 z [6, str. 108]

$$T = \mathbb{R}^+ \qquad Z = \mathbb{R}^+$$
$$\varphi(x) = x^{\frac{1}{4}} \qquad \psi(x) = x^4$$
$$\omega(x) = x^{\frac{4}{3}}$$
$$\Omega(\delta) = \textstyle\sum_{i=1}^{n} c_i \delta_i$$

Príklad 3.: cvičenie 7 z [6, str. 116]

$$T = \mathbb{R}^+ \qquad Z = \mathbb{R}^+$$
$$\varphi(x) = \frac{1}{x} \qquad \psi(x) = \frac{1}{x}$$
$$\omega(x) = x^{\frac{1}{2}}$$
$$\Omega(\delta) = \textstyle\sum_{i=1}^{n} c_i \delta_i$$

2.2 Zhrnutie modelu

V tejto kapitole sme popísali všeobecný model, ktorý zahŕňa všetky tri príklady z predchádzajúcej kapitoly. Tento model je charakterizovaný všeobecnou rýdzomonotónnou funkciou $\varphi : T \to Z$ a jej inverznou funkciou $\varphi^{-1} = \psi : Z \to T$. V nasledujúcej časti sa budeme snažiť zistiť, pre aké triedy funkcií φ bude možné skutočne zostrojiť duálnu úlohu, t.j. explicitne skonštruovať príslušné funkcie $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ a $\Omega : \mathbb{R}_+^m \to \mathbb{R}$. Počas odvodenia sme sformulovali tri netriviálne vlastnosti, ktoré bude musieť funkcia φ spĺňať:

1.vlastnosť: Funkcia $f(z) = \varphi(g(z))$ musí byť konvexná pre rastúce φ. , resp. konkávna pre klesajúce φ.

2.vlastnosť: Musí existovať funkcia ω, ktorá spĺňa (60) a (62), t.j. $\sum_{i=1}^{n} \omega[\delta_i] = 1$, kde $\delta_i := \varphi'(g(z)) \cdot \psi'(c_i + x_i)$.

3.vlastnosť: Musí sa dať zostrojiť funkcia $\Omega(\delta)$ ako výsledok nahradenia premennej z v účelovej funkcii $f(z)$ novou premennou δ zo vzťahu (60). Takto zostrojená $\Omega(\delta)$ bude duálnou účelovou funkciou vo vzťahu (57) resp. (58) .

V nasledujúcej kapitole ukážeme, že zovšeobecné funkcie zahŕňajúce príklady 1, 2, 3 spĺňajú všetky tri uvedené vlastnosti. V kapitole 4 ukážme, že tieto funkcie sú jediné, čo spĺňajú vlastnosť 2. A keďže funkcia, čo zopovedá nášmu modelu musí napĺňať všetky tri uvedené vlastnosti, zoznam riešení, ktorý bude uvedený v kapitole 3, je úplný.

3 Postačujúce podmienky funkcií vo všeob. modeli

V tejto kapitole ukážeme, že 2 mierne zovšeobecnené funkcie z príkladov 1, 2, 3, spĺňajú všetky tri vlastnosti spomenuté v závere predchádzajúcej kapitoly. Prvou z nich bude logaritmická funkcia a druhou bude mocninová funkcia.

3.1 Logaritmická funkcia

V príklade 1 sme mali funkciu $\varphi(x) = ln(x)$. Tú možno zovšeobecniť na tvar $\varphi(x) = b \cdot lnx + d$. Pre tento funkčný predpis možno explicitne vyjadriť predpisy všetkých ostatných funkcií φ', ψ, ψ', ω, ktoré sa v modeli vyskytujú:

$$\begin{aligned} \varphi(x) &= b \cdot lnx + d & \varphi'(x) &= b \cdot x^{-1} \\ \psi(x) &= e^{\frac{x-d}{b}} & \psi'(x) &= e^{\frac{x-d}{b}} \cdot \frac{1}{b} \\ & & \omega(x) = x & \end{aligned} \tag{65}$$

1.vlastnosť je pre $b > 0$ splnená analogicky ako v príklade 1, pričom dôkaz je uvedený v [6, str.110]. Ak je $b < 0$, tak sa účelová funkcia mení na konkávnu, a keďže trasformácia je klesajúca, minimalizácia funkcie $g(z)$ sa mení na maximalizáciu funkcie $f(z)$. Potom Wolfeho duálnu úlohu možno zostrojiť a bude to úloha maximalizačná.

2.vlastnosť t.j. overenie vzťahu (62) pre funkciu $\omega(x) = x$, t.j. $\sum_{i=1}^{n} \omega[\delta_i] = 1$, kde $\delta_i = \varphi'\left(\sum_{k=1}^{n} \psi(c_k + x_k)\right) \cdot \psi'(c_i + x_i)$. Podrobne rozpíšme tento vzťah :

$$\sum_{i=1}^{n} \omega \left[\varphi' \left(\sum_{k=1}^{n} \psi(c_k + x_k)\right) \cdot \psi'(c_i + x_i)\right] = 1, \tag{66}$$

Čo podľa vzťahov (65) dáva:

$$\sum_{i=1}^{n} \left[b \cdot \left(\sum_{k=1}^{n} \left(e^{\frac{c_k + x_k - d}{b}}\right)\right)^{-1} \cdot e^{\frac{c_i + x_i - d}{b}} \cdot \frac{1}{b}\right]. \tag{67}$$

Vo vzťahu (67) sa b vykráti, a keďže vnútorná suma je nezávislá od vonkajších indexov, možno ju vybrať von. Teraz už rovnosť s ľavou stranou platí triviálne:

$$\sum_{i=1}^{n} \left[e^{\frac{c_i + x_i - d}{b}}\right] \cdot \frac{1}{\sum_{k=1}^{n} \left(e^{\frac{c_k + x_k - d}{b}}\right)} = 1. \tag{68}$$

Tým sme ukázali, že vzťahy (65) spĺňajú vlastnosť 2.

3.vlastnosť Postup na zostrojenie duálnej funkcie pre $b = 1$, $d = 0$ bol podrobne urobený v kapitole 1, v príklade 1 - Geometrické programovanie. Zostrojenie funkcie s parametrami b, d bude analogické ako v spomínanom príklade. Tvar primárnej úlohy je :

$$Min\left\{ f(z) = b \cdot ln\left(\sum_{i=1}^{n} e^{\frac{x_i + c_i - d}{b}}\right) + d \,\middle|\, x_i = \sum_{j=1}^{m} a_{ij} z_j, z \in \mathbb{R}^m \right\}, \text{ kde } c_i > 0 \text{ a } a_{ij} \in \mathbb{R}. \tag{69}$$

A tvar úlohy duálnej k úlohe (69) je[5]:

$$Max\left\{ b\sum_{i=1}^{n} \delta_i ln\left(\frac{e^{\frac{c_i}{b}}}{\delta_i}\right) \,\middle|\, \sum_{i=1}^{n} \delta_i a_{ij} = 0, \text{ pre } \forall j = 1 \dots m. , \sum_{i=1}^{n} \delta_i = 1, \delta \in \mathbb{R}_+^m \right\}. \tag{70}$$

Vidno, že v duálnej úlohe sa už nenachádza nepodstatný parameter d, a taktiež parameter b neprináša nové možnosti a možno ho vynechať. V tomto zmysle parametrizácia nepriniesla podstatné zovšeobecnenie príkladu 1.

3.2 Mocninová funkcia

Budeme analyzovať funkciu $\varphi(x) = \frac{b}{c+1} \cdot x^{c+1} + d$ [6], ktorej špeciálne prípady boli analyzované v príklade 2, kde $b = \frac{1}{4}, c = -\frac{3}{4}, d = 0$ a v príklade 3, kde $b = -1, c = -2, d = 0$. Pre tento funkčný predpis možno explicitne vyjadriť predpisy všetkých ostatných funkcií φ', ψ, ψ', ω, ktoré sa v modeli vyskytujú:

$$\begin{aligned} \varphi(x) &= \frac{b}{c+1} \cdot x^{c+1} + d & \varphi'(x) &= b \cdot x^c \\ \psi(x) &= \left[\frac{x-d}{b} \cdot (c+1)\right]^{\frac{1}{c+1}} & \psi'(x) &= \left[\frac{x-d}{b} \cdot (c+1)\right]^{-\frac{c}{c+1}} \cdot \frac{1}{b} \end{aligned} \tag{71}$$

$$\omega(x) = x^{\frac{-1}{c}} \text{ , kde } c \neq -1, c \neq 0.$$

[5] Pre $b < 0$ primárna úloha (69) musí byť maximalizačná a následne duálna úloha (70) bude minimalizačná.

[6] Najdôležitejším parametrom je v tomto prípade c, ktorý výrazne mení charakter úlohy.

1.vlastnosť: Pre overenie tejto vlastnosti je nutné rozdeliť mocninové funkcie na dve vetvy určené konvexnosťou, resp. konkávnosťou účelovej funkcie :

$$f(z) = \tfrac{b}{c+1} \cdot \left[\sum_{i=1}^{n} \left([c_i + x_i - d] \cdot \tfrac{c+1}{b} \right)^{\frac{1}{c+1}} \right]^{c+1} + d.$$

Prvou vetvou sú funkcie, kde $c \in (-1; 0)$. Pre tieto hodnoty parametra c, odmysliac si koeficient $\frac{b}{c+1}$, sa účelová funkcia správa ako p-norma. Keďže každá norma je konvexnou funkciou, konvexnosť, resp. konkávnosť účelovej funkcie závisí, len od b. Ak je $b > 0$, ide o minimalizáciu konvexnej funkcie a ak je $b < 0$, ide o maximalizáciu konkávnej funkcie. Teda je splnená podmienka pre vytvorenie Wolfeho duálnej úlohy.

Druhou vetvou sú funkcie, kde $c < -1$. Pre tieto hodnoty parametra c, odmysliac si kladný koeficient $\frac{b}{c+1}$, je primárna účelová funkcia:

$$f(z) = \frac{b}{c+1} \cdot \left[\sum_{i=1}^{n} \left([c_i + x_i - d] \cdot \frac{c+1}{b} \right)^{\frac{1}{c+1}} \right]^{c+1} + d$$

konkávna (Dôkaz tohto tvrdenia sa nachádza v prílohe A). Preto konvexnosť, resp. konkávnosť účelovej funkcie, závisí len od parametra b. Ak je $b > 0$, ide o minimalizáciu konvexnej funkcie a ak je $b < 0$, ide o maximalizáciu konkávnej funkcie. Teda je splnená podmienka pre vytvorenie Wolfeho duálnej úlohy.
Treba povedať, že triedu mocninových funkcií, kde $c \in \langle 0; \infty)$ musíme z ďalšieho uvažovania vylúčiť, pretože ich vlastnosti neumožňujú zostrojiť Wolfeho duálnu úlohu.

2.vlastnosť, t.j. overenie vzťahu (62) pre funkciu ω t.j. $\sum_{i=1}^{n} \omega[\delta_i] = 1$, kde $\delta_i = \varphi'\left(\sum_{k=1}^{n} \psi(c_k + x_k)\right) \cdot \psi'(c_i + x_i)$. Zo vzťahov (71) dosadených do vzťahu (66) získavame: [7]

$$\sum_{i=1}^{n} \left[b \cdot \left(\sum_{k=1}^{n} \left[\frac{c_k + x_k - d}{b} \cdot (c+1) \right]^{\frac{1}{c+1}} \right)^{c} \cdot \left[\frac{c_i + x_i - d}{b} \cdot (c+1) \right]^{-\frac{c}{c+1}} \cdot \frac{1}{b} \right]^{-\frac{1}{c}} . \quad (72)$$

Vo vzťahu (72) sa b vykráti, a keďže vnútorná suma je nezávislá od vonkajších indexov, možno ju vybrať von. Teraz už rovnosť s ľavou stranou platí triviálne:

$$\sum_{i=1}^{n} \left[\left[\frac{c_i + x_i - d}{b} \cdot (c+1) \right]^{-\frac{c}{c+1}} \right]^{-\frac{1}{c}} \left(\sum_{k=1}^{n} \left[\frac{c_k + x_k - d}{b} \cdot (c+1) \right]^{\frac{1}{c+1}} \right)^{-1} . \quad (73)$$

[7] Toto overenie nezávisí od parametra c a prebehne pre ľubovoné $c \neq -1, c \neq 0$.

Tým sme ukázali, že vzťahy (71) spĺňajú vlastnosť 2.

3.vlastnosť Postup na zostrojenie duálnej funkcie pre $b = \frac{1}{4}, c = -\frac{3}{4}, d = 0$, kde $c \in (-1; 0)$, je podrobne urobený v kapitole 1, v príklade 2.
Postup na zostrojenie duálnej funkcie pre $b = -1, c = -2, d = 0$, kde $c < -1$, je podrobne urobený v kapitole 1, v príklade 3.
Zostrojenie funkcie s parametrami b, c, d bude analogické ako v spomínaných príkladoch. Tvar primárnej úlohy je :

$$Min\left\{\frac{b}{c+1}\left[\sum_{i=1}^{n}\left(\left[c_i + x_i - d\right]\cdot\frac{c+1}{b}\right)^{\frac{1}{c+1}}\right]^{c+1}\middle| x_i = \sum_{j=1}^{m} a_{ij}z_j, z \in \mathbb{R}_+^m\right\}. \qquad (74)$$

Tvar duálnej úlohy k úlohe (74) je[8]:

$$Max\left\{V(\delta) = \sum_{i=1}^{n}(c_i - d)\,\delta_i \middle| \sum_{i=1}^{n}\delta_i a_{ij} = 0, \text{ pre } \forall j = 1\ldots m.\ , \sum_{i=1}^{n}\delta_i^{-\frac{1}{c}} = 1, \delta \in \mathbb{R}_+^m\right\}. \qquad (75)$$

Z dvojice duálnych úloh vidno, že parametre b a d nie sú podstatným zovšeobecnením. Avšak parameter c v exponente zahŕňa optimalizačné úlohy pre ľubovlnú p-normu, ako aj pre "záporné mocninové priemery[9]".

[8]Zostrojené úlohy sú pre parameter $b > 0$, ak je $b < 0$, tak primárna úloha (74) je maximalizačná a duálna úloha (75) je minimalizačná.

[9] Ak položíme $c_i = 0$, $d = 0$, $c + 1 = \frac{1}{p}$, $b = \frac{1}{p}n^{\frac{1}{p}}$, potom dostávame účelovú funkciu v tvare $s(x) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}$, kde $p < 0$.

Uveďme ešte raz postup, akým možno využiť duálnu úlohu k vyriešeniu primárnej:

Zhrnutie:

1. Chceme riešiť úlohu (74).
2. Namiesto nej vyriešime jednoduchšiu úlohu (75). Získame $\hat{\delta} \in \mathbb{R}^m_+$ a $V\left(\hat{\delta}\right)$.
3. Aplikovaním teórie triviálnej duality priamo získavame hodnotu optima úlohy (74): $f(\hat{z}) = V\left(\hat{\delta}\right)$.
4. Použitím substitučných vzťahov: $\delta_i \left[(c_i + x_i - d)\left(\frac{c+1}{b}\right)\right]^{\frac{c}{c+1}} = \left[\sum_{k=1}^{n} \left[(c_k + x_k - d)\left(\frac{c+1}{b}\right)\right]^{\frac{1}{c+1}}\right]^c$ získame sústavu n linárnych rovníc o m neznámych $\hat{z}_j$. Z nej dopočítame hodnotu argumentu optima, teda $\hat{z} \in \mathbb{R}^m$. Sústava má tvar:

$$\sum_{j=1}^{m} a_{ij}\hat{z}_j = \hat{\delta}_i^{-\frac{c+1}{c}} \cdot \sum_{k=1}^{n} (c_k - d)\,\hat{\delta}_i - (c_i - d) \ \text{, pre } \forall i = 1 \ldots n. \tag{76}$$

4 Nutné podmienky funkcií vo všeob. modeli

4.1 Vlastnosti funkcií v modeli

V tejto kapitole ukážeme, že logaritmická a mocninová funkcia φ z kapitoly 3 sú jedinými funkciami, ktoré spĺňajú vlastnosť 2. Teda je k nim možné zostrojiť funkciu ω. Pretože dané funkcie vyhovujú všeobecnému modelu a sú jediné, čo spĺňajú vlastnosť 2, sú jediným riešením všeobecného modelu z kapitoly 2. Pripomeňme podmienky pre funkcie φ, ψ, ω, ktoré boli uvedené v kapitole 2 - Všeobecný model:

(P1): $\forall z \in Z \subset \mathbb{R} : \varphi\left(\psi\left(z\right)\right) = z$,

(P2): $\forall t \in T \subset \mathbb{R} : \psi(\varphi(t)) = t$,

(P3): φ, ψ, sú C^1 diferencovateľné,

(P4): $\omega : \mathbb{R}^+ \to \mathbb{R}^+$ je rýdzomonotónna, spojitá,

Vlastnosť 2: $\sum_{i=1}^{n} \omega\left[\delta_i\right] = 1$, kde $\delta_i = \varphi'\left(\sum_{k=1}^{n} \psi\left(c_k + x_k\right)\right) \cdot \psi'\left(c_i + x_i\right)$.

Poznámky k podmienkam na funkcie φ, resp. ψ :
V (P2) si možno lineárnou transformáciou funkcie φ zabezpečiť, že bude platiť $\langle 0; 1\rangle \subseteq T$, čo budeme ďalej využívať.
(P3): je jediná podmienka, ktorá priamo nepochádza z modelu, avšak táto vlastnosť nie je veľmi obmedzujúca a je pomerne logická.
Vo vlastnosti 2 budeme od posunu u jednotlivých premenných x_i o konštantu c_i bez ujmy na všeobecnosti abstrahovať.

4.2 Odvodenie nutných podmienok I.časť

Prepíšme si vlastnosť 2 s dosadenými premennými δ_i:

$$\sum_{i=1}^{n} \omega \left[\varphi' \left(\sum_{k=1}^{n} \psi \left(x_k \right) \right) \cdot \psi' \left(x_i \right) \right] = 1. \tag{77}$$

Vzťah (77) je vlastne diferenciálnou rovnicou 1. rádu s inverznou funkciou a divokým argumentom, pričom neznáme sú funkcie ω a ψ. Vhodnými úpravami ju prevedieme na funkcionálnu rovnicu. Treba si uvedomiť, že zafixovaním jedného alebo viacerých parametrov $n \in \mathbb{N}$, $\forall i = 1 \ldots n : x_i \in Z$, získavame nutné podmienky pre neznáme funkcie φ, ψ, ω.

Voľbou parametrov $x_i := x$ pre $\forall i = 1 \ldots n$, vzťah (77) zjednodušíme na $\sum_{i=1}^{n} \omega \left[\varphi' \left(\sum_{k=1}^{n} \psi \left(x \right) \right) \cdot \psi' \left(x \right) \right] = 1$, čo sa dá zapísať takto:

$$n\omega \left[\varphi' \left(n\psi \left(x \right) \right) \cdot \psi' \left(x \right) \right] = 1. \tag{78}$$

Pokračujme úpravou vzťahu (78) a zvoľme v ňom za $x := \varphi \left(t \right)$, čo je ekvivalentné $\psi \left(x \right) := t$, keďže funkcie φ a ψ sú si na svojich definičných oboroch inverzné. Následne možno napísať vzťah:

$$n\omega \left[\varphi' \left(nt \right) \cdot \psi' \left(\varphi \left(t \right) \right) \right] = 1. \tag{79}$$

Derivovaním (P2) dostávame obmenu známeho vzťahu pre deriváciu inverznej funkcie v tvare $\forall t \in T : \psi' \left(\varphi \left(t \right) \right) \cdot \varphi' \left(t \right) = 1$. Dosadením do rovnice (79) získame oveľa jednoduchší vzťah, ktorý už nie je diferenciánou rovnicou, ale len funkcionálnou rovnicou s neznámymi funkciami ω a φ':

$$\omega \left[\frac{\varphi' \left(nt \right)}{\varphi' \left(t \right)} \right] = \frac{1}{n}. \tag{80}$$

Podľa (P4) je ω rýdzomonotónna, a teda k nej existuje inverzná funkcia. To nám umožňuje upraviť (80) na tvar:

$$\frac{\varphi' \left(nt \right)}{\varphi' \left(t \right)} = \omega^{-1} \left[\frac{1}{n} \right]. \tag{81}$$

Pravú stranu označme ako $f \left(n \right) := \omega^{-1} \left[\frac{1}{n} \right]$. Funkcia f je zložením dvoch funkcií. Prvou z nich je $\omega^{-1} \left(x \right)$, o nej vieme, že je prostá a spojitá na svojom definičnom obore.

Druhou je $\frac{1}{x}$, táto funkcia je pre $x \in \mathbb{R}^+$ prostá a spojitá. Preto funkcia f je tiež prostá a spojitá. Po takomto označení má vzťah (81) podobu:

$$f(n) = \frac{\varphi'(nt)}{\varphi'(t)}\text{ , kde } t \in T, n \in \mathbb{N}. \tag{82}$$

Všimnime si, že voľbou $n = 1$ vo vzťahu (82) získame nutnú podmienku pre f, a to:

$$1 = f(1) = \omega^{-1}(1) = \omega(1). \tag{83}$$

Vo vzťahu (82) si možno ďalej všimnúť, že podiel $\frac{\varphi'(nt)}{\varphi'(t)}$ je vždy kladný, pretože je to podiel derivácií monotónnych funkcií. Preto o f možno povedať, že posiela každé prirodzené číslo n na kladné reálne číslo. Upravme teraz vzťah (82) do podoby:

$$\varphi'(nt) = f(n) \cdot \varphi'(t). \tag{84}$$

Zvoľme t $:= m \cdot v$, kde $m \in \mathbb{N}$ a $v \in \langle 0; 1\rangle$. Takto má náš vzťah podobu:

$$\varphi'(nmv) = f(n) \cdot \varphi'(mv). \tag{85}$$

Teraz môžeme použiť na pravú aj ľavú stranu (85) vzťah (84), takto dostávame :

$$f(nm) \cdot \varphi'(v) = \varphi'(nmv) = f(n) \cdot \varphi'(mv) = f(n) \cdot f(m) \cdot \varphi'(v). \tag{86}$$

Vydelením vzťahu (86) výrazom $\varphi'(v)$ dostávame :

$$f(nm) = f(n) \cdot f(m). \tag{87}$$

Vzťah (87) pripomína Cauchyho multiplikatívnu funkcionálnu rovnicu (Cauchy's multiplicative fuctional equation), ktorá aj s jej riešením bola publikovaná už v roku 1821 v knihe [2, str. 110]. Nasleduje spomínaná veta:

Veta 4.1. *Nech $f : \mathbb{R}^+ \to \mathbb{R}^+$ je spojitá funkcia, ktorá na množine kladných racionálnych čísel spĺňa vzťah:*

$$f(xy) = f(x) \cdot f(y). \tag{88}$$

Potom funkcia f na množine reálnych čísel musí spĺňať vzťah :

$$f(x) = x^c,\ kde\ c \in \mathbb{R}. \tag{89}$$

Naša úloha sa líši v tom, že podmienka (88) je splnená len pre všetky prirodzené čísla.
Ďalej ukážeme, že riešením v našom prípade bude $f(n) = n^c$, pre všetky prirodzené čísla n. Tento výsledok určí nutné podmienky pre funkcie :

- $\omega(x)$, ktorá je viazaná vzťahom $f(n) := \omega^{-1}\left[\frac{1}{n}\right]$,
- $\varphi'(x)$, ktorá je viazaná vo vzťahu $\varphi'(nt) = f(n) \cdot \varphi'(t)$.

Nasledovná časť bude venovaná prechodu od vzťahu (87) ku $f(n) = n^c$.

4.3 Cauchyho multiplikatívna rovnica v prirodzených číslach

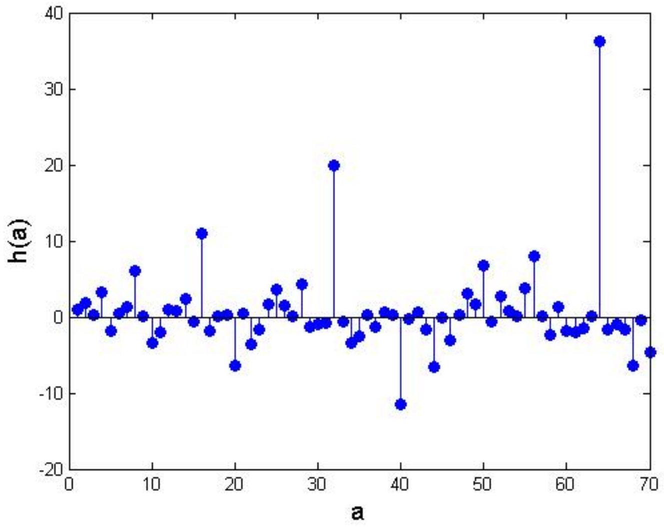
S problematikou súvisiacou s riešením Cauchyho multiplikatívnej rovnice sa môžme stretnúť v [2], [8], [9], [11], [12], [13] a v mnohých ďalších. O zúženej verzii na prirodzené čísla sa možno dočítať v [12, str. 82], kde sa píše, že riešením sú takzvané "completly multiplicative functions". Platí, že riešením je nulová funkcia a všetky funkcie, ktoré spĺňajú $f(\prod_{i=1}^{K} p_i^{\alpha_i}) = \prod_{i=1}^{K} f^{\alpha_i}(p_i)$ ak $f(1) = 1$, kde súčin predstavuje rozklad čísla na súčin provočiniteľov. Príkladmi funkcií s touto vlastnosťou sú mocninové funkcie n^c, Legeandrov symbol, či Liovillova funkcia. Ďalším príkladom je funkcia $h : A \to B, A = \langle 1; 70\rangle \cap \mathbb{N}, B \subset \mathbb{R}$ zakreslená na obrázku (1). [10]

Avšak v našom prípade máme ešte ďalšie dodatočné podmienky:

- $f(1) = 1$, zo vzťahu (83),
- $f : \mathbb{R}^+ \to \mathbb{R}^+$, zo vzťahu (82),
- $f(n) = \omega^{-1}\left[\frac{1}{n}\right]$, kde ω je z (P4) rýdzomonotónna a spojitá.

Prvý vzťah vylučuje triviálny - nulový prípad. Druhý vzťah zužuje obor hodnôt nami hľadanej funkcie. Z tretieho vzťahu vyplýva, že f je rýdzomonotónna a spojitá.

[10] Funkčné hodnoty $h(n)$ sme vypočítali tak, že pre každé prvočíslo $h(p)$ nadobúda konštantnú náhodne zvolenú hodnotu. Funkčná hodnota v zložených číslach bola dopočítaná pomocou rozkladu na súčin prvočísel a roznásobením funkčných hodnôt v nich nadobúdaných.



Obr. 1: $h : A \to B, A = \langle 1; 70 \rangle \cap \mathbb{N}, B \subset \mathbb{R}$

S využitím týchto dodatočných podmienok môžme ukázať, že jedinými vhodnými funkciami budú mocninové funkcie $f(n) = n^c$. O tom už ale bude hovoriť nasledujúca lema a veta.

Lema 4.2. *Nech $g : \mathbb{N}_0 \to \mathbb{R}$ spĺňa vzťah :*

$$\forall r, s \in \mathbb{N}_0 : g(r+s) = g(r) + g(s) . \tag{90}$$

Potom pre g platí :

$$\forall n \in \mathbb{N}_0 : g(n) = nc \text{ ,kde } c \in \mathbb{R}. \tag{91}$$

Dôkaz. Zvoľme $r = s := 0$ a dostávame $g(0) = g(0) + g(0)$, z toho $g(0) = 0$. Ďalej označme $g(1) = c$. Matematickou indukciou ukážeme, že $\forall n \in \mathbb{N}_0 : g(n) = c \cdot n$.

1. $n = 1$, vtedy $g(1) = c$, čo platí,

2. $g(n+1) = g(n) + g(1) =^{IP} cn + c = (n+1)c$, čo sme chceli dokázať.

□

Tu uvádzaná lema je riešením Cauchyho aditívnej rovnice nad prirodzenými číslami[11]. Tu uvedený dôkaz odznel na prednáške [13], ale dá sa napríklad nájsť aj v [8].

[11] Dôkaz možno spraviť aj nad racionálnymi číslami, ale pre nás to nieje potrebné.

Veta 4.3. *Nech $f : \mathbb{R}^+ \to \mathbb{R}^+$ je rýdzomonotónna funkcia, ktorá na množine prirodzených čísel spĺňa vzťah:*

$$\forall n, m \in \mathbb{N} : f(nm) = f(n) f(m). \tag{92}$$

Potom funkcia f na množine prirodzených čísel musí spĺňať vzťah:

$$\forall n \in \mathbb{N} : f(n) = n^c, \ kde\ c \in \mathbb{R} \wedge c \neq 0. \tag{93}$$

Dôkaz. Zlogaritmujme vzťah (92) logaritmom pri základe k, kde $k \in \mathbb{N} \wedge k \neq 1$. Poznamenajme, že operácia logaritmovania je korektná, pretože $f(n) > 0$. Takto získame vzťah:

$$log_k(f(nm)) = log_k(f(n)) + log_k(f(m)). \tag{94}$$

Pokračujme v úpravách a zvoľme substitúcie $n := k^r$ a $m := k^s$. Tieto substitúcie sú opäť korektné, pretože obrazom exponenciály je celé $\mathbb{R}^+$, teda $\forall n\ \exists r : k^r = n$. Získavame vzťah:

$$log_k\left(f\left(k^{r+s}\right)\right) = log_k(f(k^r)) + log_k(f(k^s)). \tag{95}$$

Takto získaný vzťah platí len pre $\forall r, s \in \{log_k(n) \mid n \in \mathbb{N}\}$. Označme túto množinu S_k. Ak si túto množinu rozpíšeme, získavame:
$S_k = \{log_k(1) = 0, log_k(2), log_k(3), ...log_k(k) = 1, ...log_k(k^2) = 2..., log_k(k^3) = 3, ...\}$.
Všimnime si, že platí inklúzia $\forall k \in \mathbb{N} : \mathbb{N}_0 \subset S_k$. Ďalej označme triedu funkcií

$$g_k(x) := log_k(f(k^x)). \tag{96}$$

S označením (96) môžme prepísať vzťah (95) ako

$$\forall r, s \in \mathbb{N}_0 : g_k(r+s) = g_k(r) + g_k(s) \ \text{, kde } k \in N \wedge k \neq 1. \tag{97}$$

Vzťah (97), spĺňa pre $\forall k \in \mathbb{N}$, $k \neq 1$ podmienky lemy (4.2). Jej aplikáciou získavame:

$$\forall n \in \mathbb{N}_0 : g_k(n) = c_k n, \tag{98}$$

kde $k \in \mathbb{N} \wedge k \neq 1$ a c_k je pre každé k ľubovolná reálna konštanta.

Vzťah (98) prevedieme pomocou (96) na funkciu f. Platí:

$$\forall n \in \mathbb{N}_0 : nc_k = log_k \left(f\left(k^n\right)\right). \tag{99}$$

Odlogaritmovaním (99) získavame $f\left(k^n\right) = \left(k^n\right)^{c_k}$. To znamená, že f sa správa ako mocninová funkcia pre každú postupnosť mocnín ľubovoného prirodzeného čísla. Ďalej ukážeme, že konštanta c_k pre každé $k \in \mathbb{N}$ okrem 1 rovnaká. Ak by nebola rovnaká pre všetky prirodzené čísla, viedlo by to k sporu s monotónnosťou funkcie f. Túto časť tvrdenia ukážeme sporom. Nech platí nasledovné: $\exists k, l \in \mathbb{N} : 1 < k < l \wedge 0 < c_k < c_l$. Funkcia f potom musí byť ostro rastúca. Ukážeme, že možno zvoliť čísla $m, n \in \mathbb{N}$, tak aby spĺňali $l^n < k^m$ a zároveň $f\left(l^n\right) > f\left(k^m\right)$. Ak také m, n budú existovať, nastáva spor s monotonnosťou f. Musia platiť vzťahy:

$$\begin{aligned} l^n &< k^m, \\ l^{nc_l} &> k^{mc_m}. \end{aligned} \tag{100}$$

Zlogaritmovaním rovníc (100) dostávame

$$\begin{aligned} n &< log_l\left(k\right) m, \\ nc_l &> log_l\left(k\right) mc_k. \end{aligned} \tag{101}$$

Oba vzťahy možno zreťaziť do vzťahu:

$$\frac{c_k}{c_l} log_l\left(k\right) < \frac{n}{m} < log_l\left(k\right). \tag{102}$$

Platí, že $0 < \frac{c_k}{c_l} < 1$ a $log_l\left(k\right) > 0$. To znamená, že v danom intervale $\exists r \in \mathbb{Q}^+$. Podielom $\frac{n}{m}$ možno vyjadriť ľubovolné kladné racionálne číslo. Zvoľme $\frac{n}{m} := r$. Nastáva spor s predpokladom monotónnosti. Teda možno povedať, že $\forall k, l \in \mathbb{N}, 1 < k < l : c_k \geq c_l$.

Analogicky budeme postupovať opačným smerom a opäť dostame spor. Nech $\exists k, l \in \mathbb{N} : 1 < k < l, \wedge c_k > c_l > 0$. Potom platí, že funkcia f musí byť ostro rastúca.Teda stačí ukázať, že $\exists m, n \in \mathbb{N}$, čo spĺňajú $k^m < l^n \wedge f\left(k^m\right) > f\left(l^n\right)$. Zopakovaním postupu z minula získavame :

$$\frac{c_l}{c_k} log_k(l) < \frac{m}{n} < log_k(l). \quad (103)$$

Pričom platí, že $0 < \frac{c_l}{c_k} < 1$ a $log_k(l) > 1$. Teda opäť možno zvoliť racionálne číslo $\frac{m}{n}$ tak, aby sme dostali spor s predpokladom monotónnosti. Preto platí $\forall k, l \in \mathbb{N}, k < l, c_k \leq c_l$.

Ďalej si možno uvedomiť, že prípad, kedy platí $c_k < 0 < c_l$ okamžite vedie k sporu s monotónnosťou a prípady, kedy platí $c_k < c_l < 0 \vee 0 > c_k > c_l$, sú analogické dvom, čo sme poriadne dokázali. Ak teda zhrnieme získané vzťahy pre c_k, c_l, získame: $\forall k, l \in \mathbb{N} : c_k = c_l$. Riešenie funkcionálnej rovnice má tvar:

$$\forall n \in \mathbb{N} : f(n) = n^c \text{ , kde } c \in \mathbb{R} \text{ a } c \neq 0. \quad (104)$$

$\square$

4.4 Odvodenie nutných podmienok II.časť

Pokračujme ďalej teraz už s explicitným predpisom pre f, ktorý sme získali vďaka vete (4.3). Dosaďme teda do vzťahu (84), ktorý viaže f a φ', predpis pre f zo vzťahu (104):

$$\forall n \in \mathbb{N}, t \in T : \varphi'(nt) = n^c \varphi'(t). \quad (105)$$

V (105) zvolíme $t := 1$, získavame tak vzťah, ktorý hovorí, že pre každé prirodzené číslo sa φ' musí správať ako násobok mocninovej funkcie:

$$\varphi'(n) = n^c \varphi'(1). \quad (106)$$

Zo vzťahu (105) možno ešte zvolením $t := \frac{m}{n}$ a $n := n$, tak aby $m, n \in \mathbb{N}$, získať nasledovné:

$$\varphi'(m) = n^c \varphi'\left(\frac{m}{n}\right). \quad (107)$$

Na vzťah (107) použijeme vzťah (106) v premennej m a upravíme ho do podoby:

$$\varphi'\left(\frac{m}{n}\right) = \left(\frac{m}{n}\right)^c \varphi'(1). \quad (108)$$

Vo vzťahu (108) zvoľme za $\frac{m}{n} := r$, potom $r \in \mathbb{Q}^+ \cap T$. Ďalej zvoľme $\varphi'(1) = b$ a náš vzťah získa podobu $\forall r \in \mathbb{Q}^+ \cap T : \varphi'(r) = br^c$. Vzhľadom k tomu, že od φ' očakávame spojitosť, možno ľahko urobiť limitný prechod k reálnym číslam, pretože k ľubovolnému reálnemu číslu možno zostrojiť konvergentnú postupnosť čísel racionálnych. Následne môžme zapísať nutnú podmienku pre φ' v tvare:

$$\forall r \in \mathbb{R}^+ \cap T : \varphi'(r) = br^c, \text{ kde b, c} \in \mathbb{R}. \tag{109}$$

K podmienke na funkciu φ sa dostaneme integrovaním vzťahu (109). Integrovanie ale musíme rozdeliť na dve vetvy, a to $c = -1$ a vetvu $c \neq -1$.

Prvá vetva:

$$\begin{aligned} c &= -1 \\ \int \varphi' dx &= \int bx^{-1} dx \\ \varphi(x) &= bln(x) + d, \end{aligned} \tag{110}$$

kde d je integračná konštanta a platí: b, d $\in \mathbb{R}$.

Druhá vetva:

$$\begin{aligned} c &\neq -1 \\ \int \varphi' dx &= \int bx^c dx \\ \varphi(x) &= \frac{b}{c+1} x^{c+1} + d, \end{aligned} \tag{111}$$

kde d je integračná konštanta a platí: b, c, d $\in \mathbb{R}$.

Vzťahy (110) , (111) jasne vymedzujú nutné podmienky pre funkciu φ. Prvá vetva určuje Geometrické programovanie a druhá mocninové funkcie. Ďalej určíme tvar funkcie ω pre obe vetvy.

Nutnú podmienku pre ω môžme získať, ak sa pozrieme na vzťah (80) a nutnú podmienku pre funkciu φ' zo vzťahu (109). Kombináciou dvoch uvedených získvame:

$$\forall t \in \mathbb{R}^+ \cap T, \forall n \in \mathbb{N} : \omega\left(\frac{b(nt)^c}{bt^c}\right) = \omega(n^c) = \frac{1}{n}. \tag{112}$$

Pravú stranu možno prepísať do tvaru $\frac{1}{n} = (n^c)^{\frac{-1}{c}}$. To nás oprávňuje k nasledujúcej formulácii nutnej podmienky:

$$\forall x \in H_c : \omega(x) = x^{\frac{-1}{c}}, \text{ kde } H_c = \{x \mid x = n^c, n \in \mathbb{N}\}. \tag{113}$$

Vzťah (113), predstavuje nutnú podmienku, ktorú musí funkcia ω spĺňať. Táto podmienka nie je tak silná ako pre funkciu φ ale predstavuje dobrý návod pre voľbu funkcie ω.

Zhrnutie: Ukázali sme, že postačujúce podmienky pre funkcie φ, resp. ψ, z kapitoly 3 sú totožné[12] ako nutné podmienky odvodené v tejto kapitole. To znamená, že sme plne popísali množinu funkcií vyhovujúcich všeobecnému modelu v kapitole 2. Okrem toho sme v závislosti od parametra c určili aj tvar funkcie ω pre jednotlivé funkcie φ, resp. ψ.

[12] Až na triedu mocninových funkcií s parametrom c>0.

Záver

Teória duality konvexného programovania je užitočný nástroj praktického riešenia optimalizačných úloh, pretože v mnohých prípadoch je duálna úloha oproti primárnej jednoduchšia.

Názorným príkladom jednoduchosti duálnej úlohy je geometrické programovanie. Geometrické programovanie je tiež výnimočné tým, že umožňuje zostrojiť duálnu úlohu k optimalizačnej úlohe bez ohraničení. Tento fakt bol dôvodom k zavedeniu pojmu triviálna dvojica duálnych úloh konvexného programovania. V [6] sa uvádzajú dva analogické príklady triviálnej duality.

V našej práci sme sa zamerali na zovšeobecnenie uvedených príkladov. V kapitole 1 sme urobili podrobný rozbor troch optimalizačných úloh - geometrického programovania bez ohraničení, minimalizácie p-normy, pre $p = 4$ a minimalizácie harmonického priemeru. Ďalej sme urobili prechod od primárnej úlohy k duálnej a popísali sme, ako zo získaného riešenia duálnej úlohy získať riešenie primárnej úlohy. Tieto príklady sme v kapitole 2 spojili pod jeden model, ktorý vytvoril priestor na hľadanie nových tried optimalizačných úloh. V kapitole 3 sme ukázali, že príklad 2 a príklad 3 sa dajú rozšíriť pre rôzne mocninové funkcie. V poslednej kapitole sme ukázali, že toto rozšírenie je maximálne, čo spadá pod všeobecný model z kapitoly 2.

Ciele práce sme naplnili popísaním príkladov využitia Wolfeho duality na zjednodušenie optimalizačnej úlohy. Príklady sme pozorne rozobrali a navrhli sme model, ktorý ich spája a podarilo sa nám určiť všetky jeho riešenia. Naše výsledky, čo sa týka úlohy bez ohraničení, sa dajú porovnať s [3], kde sa v závere ukáže, ako odvodiť duálne úlohy pre p-normy pomocou zovšeobecnenej nerovnosti medzi aritmetickým a geometrickým priemerom.

Táto práca ma obohatila v mnohých smeroch. Zlepšil som sa v práci s LaTeXom, ako aj s Matlabom. Pokrok u mňa nastal i po jazykovej stránke, v t.j. slovenskej štylistike a anglickej slovnej zásobe. Taktiež sa štúdiom literatúry výrazne prehĺbili moje znalosti duality nelineárneho programovania, schopnosti riešenia funkcionálnych rovníc, ako aj vedomosti o konvexných funkciách.

Zoznam použitej literatúry

[1] Avriel, M.: *Nonlinear programming analysis and methods*, Dover publications, New York, 1976

[2] Cauchy, L. A.: *Cours d'analyse de l'École royale polytechnique*, Imprimerie royale, Paris, 1821, dostupné na internete (17.5.2013):
http://archive.org/details/coursdanalysede00caucgoog

[3] Duffin, L. J., Peterson, E. L., Zener, C.: *Geometric Programming - Theory and Application*, John Wiley & sons, New York, 1967

[4] Castilo, E., Iglesias, A., Ruíz-Cobo, R.: *Functional Equations in Applied Sciences*, Elsevier, Amsterdam, 2005

[5] Hamala, M.: *Nelineárne programovanie*, prednášky, FMFI UK, Bratislava 2012

[6] Hamala, M.: *Nelineárne programovanie*, Vydavateľstvo technickej a ekonomickej literatúry, Bratislava, 1972

[7] Hamala, M., Trnovská, M.: *Nelineárne programovanie, teória a algoritmy*, EPOS, Bratislava, 2013

[8] Jakobsen, S. K.: *Cauchy's functional equation*, blog, 2010, dostupné na internete (2.3.2013):
http://sunejakobsen.files.wordpress.com/2010/12/cauchy_eng4.pdf

[9] Kannappan, P.: *Functional Equations an Inequalities with Applications*, Springer, New York 2009

[10] Kuczma, M.: *An Introduction to the theory of Functional Equations and Inequalities, Cauchy's Equation and Jensen's Inequality*, Birkhäuser, Berlin 2009

[11] Radovanovic, M.: *Functional equations*, Olympiad training materials, IMO compendium group, 2007, dostupné na internete (6.12.2012):
http://www-bcf.usc.edu/ lototsky/PiMuEp/FunctionalEquations.pdf

[12] Small, Ch.S.: *Functional Equations and How to Solve Them*, Springer, New York 2007

[13] Ševčovič, D.: *Matematická analýza 4*, prednášky, FMFI UK, Bratislava 2011

Príloha A (Dôkaz tvrdenia zo str. 17 a 27)

Veta A.4

Funkcia $f : \mathbb{R}_+^n \to \mathbb{R}_+$ *s predpisom :*

$$f(x) = \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}}$$

je pre $p < 0$ *konkávna.*

Dôkaz. Počítajme druhé parciálne derivácie funkcie f. Nediagonálne členy Hessovej matice majú tvar:

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \left(\frac{1}{p} - 1\right) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-2} \cdot x_j^{p-1} \cdot x_k^{p-1}.$$

Diagonálne členy Hessovej matice majú tvar:

$$\frac{\partial^2 f}{\partial x_j^2} = \left(\frac{1}{p} - 1\right) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-2} \cdot x_j^{p-1} \cdot x_j^{p-1} + (p-1) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-1} \cdot x_j^{p-2}.$$

Označme vektor zložený, z x_i^{p-1} ako x^{p-1}, a vektor zložený, z x_i^{p-2} ako x^{p-2}. Teraz možno Hessovu maticu H zapísať v tvare :

$$H = \left(\frac{1}{p} - 1\right) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-2} \left(x^{p-1}\right) \left(x^{p-1}\right)^T + (p-1) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-1} \cdot diag\left(x^{p-2}\right).$$

Podmienkou pre konkávnosť je, aby jej Hessova matica funkcie bola záporne semidefintná. Matica H je súčtnom 2 matíc, prvá z nich je:

$$H_1 = \left(\frac{1}{p} - 1\right) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-2} \left(x^{p-1}\right) \left(x^{p-1}\right)^T.$$

Táto matica je záporne semidefinitná, pretože $\left(x^{p-1}\right)\left(x^{p-1}\right)^T$ je kladne definitná matica, konštanta pred ňou je záporná. Druhá matica je:

$$H_2 = (p-1) \cdot \left[\sum_{i=1}^{n} x_i^p\right]^{\frac{1}{p}-1} \cdot diag\left(x^{p-2}\right).$$

Táto matica je záporne definitná, a to preto, že je to záporným číslom prenásobená kladne definitná matica $diag\left(x^{p-2}\right)$. Hessova matica H je teda súčtom záporne definitnej a záporne semidefinitnej matice, teda je záporne definitná.

$\square$