Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Marek Olos

Konvexita v úlohách s pravděpodobnostními omezeními

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	RNDr. Ing. Miloš Kopa Ph.D.
Studijní program:	Matematika
Studijní obor:	obecná matematika

Praha 2013

Týmto chcem poďakovať RNDr. Ing. Milošovi Kopovi Ph.D. za cenné konzultácie, ochotu a zapožičanie softwaru a literatúry, rodine za podporu.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

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 v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne .... Podpis autora

Název práce: Konvexita v úlohách s pravděpodobnostními omezeními

Autor: Marek Olos

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Ing. Miloš Kopa Ph.D.

Abstrakt: Tato práce se zaměřuje na úlohy stochastického programování s pravděpodobnostními omezeními. První kapitola je úvod. Ve druhé kapitole formulujeme několik úloh stochastického programování. Ve třetí kapitole předkládáme teorii α-konkávních funkcí a mír jako základní nástroj pro vyšetřování konvexity úloh a formulujeme postačující podmínky pro konvexitu úloh zavedených v kapitole 2 pro spojité rozdělení náhodných vektorů. Důsledky teorie pak použijeme na charakterizování velké třídy spojitých rozdělení splňujících postačující podmínky pro konvexitu a na dokazování konvexity konkrétních množin. Ve čtvrté kapitole předkládáme postačující podmínky pro konvexitu úloh pro diskrétní rozdělení a v krátkosti se věnujeme metodě p-level eficientní bodů. V páté kapitole řešíme úlohu optimalizace portfolia pomocí Kataokovho modelu.

Klíčová slova: stochastické programování, pravděpodobnostní omezení, konvexita

Title: Convexity in chance constraints programming

Author: Marek Olos

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Ing. Miloš Kopa Ph.D.

Abstract: This thesis deals with chance constrained stochastic programming problems. The first chapter is an introduction. We formulate several stochastic programming problems in the second chapter. In chapter 3 we present the theory of α-concave functions and measures as a basic tool for proving convexity of the problems formulated in chapter 2 for the continuous distributions of the random vectors. We use the results of the theory to characterize a large class of the continuous distributions, that satisfy the sufficient conditions for the convexity and to prove convexity of concrete sets. In chapter 4 we present sufficient conditions for the convexity of the problems and we briefly discuss the method of the p-level efficient points. In chapter 5 we solve a portfolio selection problem using Kataoka's model.

Keywords: stochastic programming, chance constraint, convexity

Názov práce: Konvexita v úlohách s pravdepodobnostnými obmedzeniami

Autor: Marek Olos

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: RNDr. Ing. Miloš Kopa Ph.D.

Abstrakt: Táto práca sa zameriava na úlohy stochastického programovania s pravdepodobnostnými obmedzeniami. Prvá kapitola je úvod. V druhej kapitole formulujeme niekoľko úloh stochastického programovania. V tretej kapitole predkladáme teóriu α-konkávnych funkcií a mier ako základný nástroj na vyšetrovanie konvexity úloh a formulujeme postačujúce podmienky pre konvexitu úloh zavedených v kapitole 2 pre spojité rozdelenia náhodných vektorov. Dôsledky teórie potom použijeme na charakterizovanie veľkej triedy spojitých rozdelení spĺňajúcich postačujúce podmienky pre konvexitu a na dokazovanie konvexity konkrétnych množín. V štvrtej kapitole predkladáme postačujúce podmienky pre konvexitu úloh pre diskrétne rozdelenia a v krátkosti sa venujeme metóde p-level eficientných bodov. V piatej kapitole riešime úlohu optimalizácie portfólia pomocou Kataokovho modelu.

Kľúčové slová: stochastické programovanie, pravdepodobnostné obmedzenia, konvexita

Kapitola 1

Úvod

Úlohy s pravdepodobnostnými obmedzeniami sú jedným zo základných typov úloh stochastického programovania. Využitie modelov s pravdepodobnostnými obmedzeniami môžeme nájsť všade, kde musíme naše správanie optimalizovať pri náhodných parametroch od telekomunikácii cez medicínu až k financiám. Nástroj pravdepodobnostného programovania napríklad pomohol zlepšiť spoľahlivosť regulácie výšky hladiny jazera Balaton v Maďarsku a udržať ju do predpísaných limitov z pôvodných 80% na $97,5\%$. Pravdepodobnostné obmedzenia sa často využívajú vo formuláciach nových modelov a to vedie k rastúcej potrebe analýzy vlastností úloh s pravdepodobnostnými obmedzeniami. Jednou z najdôležitejších a najpreskúmanejších vlastností optimalizačných úloh je ich tzv. konvexita, ktorej sa budeme venovať v tejto práci.

Cieľom našej práce bolo zformulovať niektoré základné úlohy s pravdepodobnostnými obmedzeniami a zamerať sa na postačujúce podmienky pre ich konvexitu, ktorá úlohu robí riešitelnou aj pre veľký počet parametrov. V druhej kapitole si vysvetlíme význam úlohy pravdepodobnostného programovania, zformulujeme si niektoré základné modely a uvedieme príklady ich využitia. V tretej kapitole prezentujeme postačujúce podmienky pre konvexitu úloh so spojitými rozdeleniami pre náhodný vektor vystupujúci v pravdepodobnostných obmedzeniach. Predstavíme si teóriu α-konkávnych funkcií a mier, ktorá tvorí základný nástroj pre vyšetrovanie konvexity úloh a množín. Ukáže sa, že konvexita úlohy sa dá vyvodiť na základe rozdelenia náhodného vektora vystupujúceho v danom modeli. Pre spojité rozdelenia bude situácia jasnejšia a v ďalších častiach druhej kapitoly si pre niektoré spojité rozdelenia, na základe ich hustoty, alebo distribučnej funkcie dokážeme konvexitu daných modelov, keď náhodný vektor má dané rozdelenie. Na základe teórie potom ukážeme konvexnosť konkrétnych množín zadaných pravdepodobnostnými obmedzeniami v sekcii riešených príkladov. Štvrtú kapitolu venujeme problematike diskrétnych rozdelení. Aj pre nich existujú postačujúce podmienky pre konvexitu úlohy, žiaľ nepokryjú toľko praktických úloh ako v prípade spojitých rozdelení. Pri riešení úlohy s diskrétnym rozdelením sa teda nemôžeme na konvexitu spoliehať, preto v práci navrhneme riešiť úlohu pomocou “p-level” eficientných bodov. V piatej kapitole si ukážeme možné využitie pravdepodobnostného programovania v praxi.

Kapitola 2

Úlohy s pravdepodobnostnými obmedzeniami

2.1 Základná úloha nelineárneho programovania

Pripomeňme si základnú úlohu nelineárneho programovania v tvare:

$$\begin{array}{c} \text{Min} \quad h(x) \\ \text{za podmienok} \\ g_i(x) \leq 0, \quad i = 1, ..., m, \end{array} \tag{2.1}$$

kde $x \in R^n$ je rozhodovací vektor, $n \in N$, $m \in N$, h, g_i sú reálne funkcie, obecne nelineárne, $i = 1, ..., m$.

Poznámka 2.1. *Pokiaľ nebude povedané inak, v ďalších úlohách $x \in R^n$ bude hľadaný rozhodovací vektor.*

Dôležitým špeciálnym prípadom úlohy (2.1) je úloha lineárneho programovania:

$$\begin{array}{c} \text{Min} \quad c^T x \\ \text{za podmienok} \\ Ax \geq b \\ x \geq 0, \end{array} \tag{2.2}$$

kde $x \in R^n$, $n \in N$, $m \in N$, A je matica $m \times n$.

V tejto práci budeme uvažovať modifikácie úloh (2.1) a (2.2) zahrňujúce náhodu.

Uvedieme si niektoré základné definície.[1]

Definícia 2.2. *Povieme, že množina $A \subseteq R^n$ je konvexná, keď pre každé dva body $x, y \in A$ a $0 < \lambda < 1$ platí $\lambda x + (1 - \lambda) y \in A$.*

Definícia 2.3. *Funkcia $f : R^n \to R$ je konvexná na množine M práve vtedy, keď M je konvexná množina a pre každé $x, y \in M$ a $0 < \lambda < 1$ platí*

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

Definícia 2.4. *Povieme, že úloha (2.1) je konvexná, keď jej množina prípustných riešení je konvexná množina a funkcia $h(\cdot)$ je konvexná na tejto množine.*

Tvrdenie 2.5. *Nech $F : R^m \to R$ je neklesajúca konvexná funkcia a $g : R^n \to R^m$ je také zobrazenie, že každá jeho zložka g_i je konvexná funkcia. Potom $F \circ g$ je konvexná funkcia.*

Viac o základných definíciách a vetách ohľadom konvexity množín a funkcií sa dá nájsť v ([1] str. 9-32).

2.2 Základný tvar úlohy s pravdepodobnostnými obmedzeniami

Základná úloha nelineárneho programovania (2.1) v sebe môže ukrývať parametre, ktoré sú náhodnými veličinami. Ignorovanie faktu, že veličina môže nadobúdať náhodné hodnoty, môže viesť k nepresným až nesprávnym výsledkom. Úloha stochastického programovania je často formulovaná ako základný deterministický problém(napr. ako úloha (2.1) alebo (2.2)), kde si uvedomíme, že niektoré parametre nie sú konštanty ale náhodné veličiny a formulácia problému je takto nepresná. Alternatívu predstavuje úloha s pravdepodobnostnými obmedzeniami:

$$\begin{gathered} \text{Min} \quad h(x) \\ \text{za podmienok} \\ h_0(x) = P(g_1(x, Z) \geq 0, ..., g_r(x, Z) \geq 0) \geq p_0 \\ g_{r+1}(x) \geq p_{r+1}, ..., g_m(x) \geq p_m, \end{gathered} \tag{2.3}$$

kde Z je náhodný vektor na pravdepodobnostnom priestore (Ω, A, P), $h(x), h_0(x)$, $g_1(x, Z), ..., g_r(x, Z), g_{r+1}(x), ..., g_m(x)$ sú dané funkcie, $0 < p_0 \leq 1$, $p_{r+1}, ..., p_m$ sú dané reálne čísla.

V prípade, že v modeli očakávame, že má prebehnúť iba jedno rozhodnutie o rozhodovacom vektore v celom výpočte vravíme, že model je statický. Ak v modeli očakávame, že rozhodnutia o rozhodovacom vektore budú príjimané tak, že medzi dvoma rozhodnutiami prebehne pozorovanie skutočnej hodnoty náhodného vektora Z, potom vravíme, že model je dynamický. V tejto práci si predstavíme a vyšetríme konvexitu iba pre statické modely. (Hoci pre niektoré dynamické modely by sa ich konvexita dala odvodiť z tu prezentovanej teórie, podrobnejšie sa o nich dá dočítať napr. v [3] str.(27-84).)
Ukážeme si niektoré prístupy, ktoré vedú k úlohe (2.3) alebo k nejakému špeciálnemu prípadu úlohy (2.3) a problémy, ktoré sa ani nedajú vyriešiť inak ako nástrojom pravdepodobnostného programovania. Dobrým príkladom, na ktorom si ukážeme ako naivný prístup môže viesť k nesprávnym výsledkom je úloha o optimalizovaní kapacity 2 rezervoárov. (Príklad je riešený v [2] str. 6-9.)

Príklad 2.6. *Uvažujeme systém dvoch rezervoárov, ktoré majú chrániť oblasť pred povodňami. Systém s rezervoármi aj s prítokmi sú zobrazené na (obr.2.1). Povodeň môže byť spôsobená dvoma náhodnými prítokmi Z_1, Z_2. To nám napovedá, že Z_1, Z_2 nebudú konštantami, ale náhodnými veličinami. Nebezpečenstvo povodne sa objaví raz za rok a prítoky prichádzajú v tom istom čase. Škody spôsobené povodňou o veľkosti $y \geq 0$ sú modelované ako konvexná neklesajúca funkcia $L(y)$, kde $L(0) = 0$. Našou úlohou je určiť kapacity rezervoárov x_1, x_2 tak, aby stredná hodnota škôd bola menšia ako určený limit b, a aby náklady na vybudovanie rezervoárov $f(x_1, x_2)$ boli čo najmenšie. Veľkosť povodne je:*

$$y = \max(0, Z_1 + Z_2 - x_1 - x_2, Z_2 - x_2)$$

Náš problém nadobúda formy:

$$\begin{array}{c} \text{Min} \quad f(x_1, x_2) \\ \text{za podmienok} \\ \mathrm{E}[L(\max\{0, Z_1 + Z_2 - x_1 - x_2, Z_2 - x_2\})] \leq b \\ x_1 \geq 0, x_2 \geq 0. \end{array} \tag{2.4}$$

Bolo by chybou nahradiť Z_1, Z_2 ich strednými hodnotami $\mu_1 := \mathrm{E}(Z_1)$, $\mu_2 := \mathrm{E}(Z_2)$ a mať tak úlohu v jednoduchšom (deterministickom) tvare. Podľa Jensenovej nerovnosti máme:

$$\begin{array}{c} L(\max\{0, \mu_1 + \mu_2 - x_1 - x_2, \mu_2 - x_2\}) \\ \leq \mathrm{E}[L(\max\{0, Z_1 + Z_2 - x_1 - x_2, Z_2 - x_2\})]. \end{array}$$

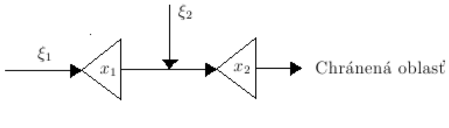
A rozdiel by mohol byť veľký dokonca aj pre lineárnu funkciu $L(y)$. Vo výsledku očakávané škody spôsobené povodňami môžu byť oveľa väčšie ako by predpovedal takýto naivný deterministický model.
Možno ešte lepšou alternatívou by bolo naviac požadovať, aby škody spôsobené povodňou neprekročili istú hranicu $d > 0$ aspoň s pravdepodobnosťou $p \in (0, 1)$. K obmedzeniam by sme naviac pridali pravdepodobnostné obmedzenie

$$P(L(y) \leq d) \geq p$$

a dostali by sme tak úlohu typu (2.3).

2.3 Náhoda v lineárnej úlohe

Veľmi často riešime prípad, že sa náhoda objaví lineárnej úlohe (2.2) Obecne môže náhoda do lineárnej úlohy vstúpiť mnohými spôsobmi. Náhoda sa môže objaviť v



Obr. 2.1: Prítoky

účelovej funkcii, pričom ostatné premenné (v obmedzeniach) sú nenáhodné. Jedná sa väčšinou o minimalizáciu (alebo maximalizáciu) strednej hodnoty účelovej funkcie. Jednoduchým príkladom môže byť maximalizácia výnosu portfólia, kde i-tá zložka náhodného vektora c je náhodným výnosom i-tej možnosti, x_i by bol objem investície do i-tej možnosti $i = 1, ..., n$. $Ax \leq b$ by bolo obmedzenie objemu investície, koľko môžeme do daných možností investovať.

Oveľa zaujímavejší prípad je, keď sa náhoda objaví (aj) v obmedzeniach.

$$\begin{gathered} \text{Min} \quad c^T x \\ \text{za podmienok} \\ \mathrm{P}(Tx \geq Z) \geq p \\ Ax \geq b, x \geq 0, \end{gathered} \tag{2.5}$$

kde $x \in R^n$ je rozhodovací vektor, $c \in R^n$ konštantný vektor $n \in N$, $m \in N$, A je matica $m \times n$, T je matica $r \times n$ $0 < p \leq 1$. Kde náhodná môže byť pravá strana nerovnosti $Tx \geq Z$ alebo matica T, prípadne oboje.

To, že často je nástroj pravdepodobnostného programovania jedinou možnosťou ako riešiť daný problém si ukážeme na nasledujúcom príklade, kde si zároveň ukážeme ako sa dá dospieť k úlohe typu (2.5), kde náhodný bude len vektor Z. Uvažujme problém nemocníc, ktoré chcú zakúpiť lôžka pre pacientov. Množstvo pacientov, ktorých je potrebné hospitalizovať v daný deň nech je Z. Našou úlohou je zakúpiť lôžka pre nemocnice, aby pacienti, ktorí prídu do nemonice mali kde byť hospitalizovaní a zároveň minimalizovať náklady na zakúpenie lôžok. Dostávame tak úlohu lineárneho programovania (2.2).
Teoreticky je možné, aby v daný deň prišli do nemocnice všetci obyvatelia mesta. Splniť takéto obmedzenie aj pre tento medzný prípad by bolo finančne nemožné. Nemôžeme tak splniť obmedzenia s pravdepodobnosťou 1. Preto je rozumnejšie riešiť radšej úlohu s obmedzeniami, kde požadujeme, aby bolo pre pacientov v daný deň pripravených dostatok lôžok len s istou pravdepodobnosťou p. Tým dostávame úlohu pravdepodobnostného programovania (2.5), kde náhodná je pravá strana nerovnosti $Tx \geq Z$ a pravdepodobnostné obmedzenie má tvar $\mathrm{P}(Tx \geq Z) \geq p$.

Jeden z najzaujímavejších prípadov náhodnej matice je spojený s Kataokovým modelom ([2] str. 284)

$$\begin{gathered} \text{Max} \quad d \\ \text{za podmienok} \\ \mathrm{P}(\sum_{i=1}^n Z_i x_i \geq d) \geq p \\ \sum_{i=1}^n x_i = M, x \geq 0, \end{gathered} \tag{2.6}$$

kde predpokladáme, že vektor $(Z_1, ..., Z_n)$ má n-rozmerné normálne rozdelenie. $0 < p \leq 1$, $M \in R$ je konštanta a $(x_1, ..., x_n, d)$ je rozhodovací vektor.
Využitie Kataokovho modelu môžeme nájsť napríklad vo financiách. M by značilo množstvo celkového investovaného kapitálu, Z_i náhodný výnos i-tej možnosti na jednotku investovaného kapitálu. Hľadané d by bola maximálna hranica, taká, že s pravdepodobnosťou aspoň p by výnos celého portfólia bol väčší alebo rovný d.

Podmienka $x \geq 0$ by zakazovala krátke pozície. Kataokov model bude ďalej analyzovaný v kapitole 3 a 4 a ilustrovaný v kapitole 5, kde pomocou neho vyriešime úlohu o optimalizovaní portfólia.

2.4 Prístupy k porušeniu obmedzení

Dobrou metódou ako nájsť správne riešenie pre danú úlohu môže byť popri pravdepodobnostných obmedzeniach penalizovať účelovú funkciu podľa miery ako prekročíme obmedzenia.([2] str. 270) Takýto hybridný model vyzerá nasledovne:

$$\begin{array}{c} \mathrm{Min}(c^T x + \sum_{i=1}^{r} q_i E([Z_i - T_i x]_+)) \\ \text{za podmienok} \\ \mathrm{P}(Tx \geq Z) \geq p \\ Ax \geq b, x \geq 0, \end{array} \tag{2.7}$$

kde $T_1, ..., T_r$ sú riadky matice T, $Z_1, ...Z_r$ sú zložky náhodného vektora Z a $q_i, ...q_r$ sú nezáporné konštanty, penalizujúce prekročenia obmedzení $T_1 x \geq Z_1, ..., T_r x \geq Z_r$.

V spojitosti s mierou porušení obmedzení sme v (2.7) penalizovali účelovú funkciu v zmysle: $\mathrm{E}([Z_i - T_i x]_+)$ je priemerne prekročené i-té obmedzenie, q_i sú váhy podľa ktorých penalizujeme účelovú funkciu pre porušenie i-tého obmedzenia.

Existujú ďalšie možnosti ako formulovať model v súvislosti s prekročením obmedzení. Ukážeme si ešte jednu z nich. Dodatočné obmedzenie v (2.8) nám pomáha vyhnúť sa možnosti príliš veľkého nesplnenia obmedzení dodatočným obmedzením: ak prekročíme obmedzenia, tak nech priemerné prekročenie obmedzení nie je priveľké.

$$\begin{array}{c} \mathrm{Min}(c^T x + \sum_{i=1}^{r} q_i E([Z_i - T_i x]_+)) \\ \text{za podmienok} \\ \mathrm{P}(Tx \geq Z) \geq p \\ \mathrm{E}(Z_i - T_i x | Z_i - T_i x > 0) \leq d_i, \quad i = 1, ..., r \\ Ax \geq b, x \geq 0, \end{array} \tag{2.8}$$

kde $d_i, i = 1, ..., r$ sú dané konštanty. Pridané obmedzenia nazývame obmedzenia podmienenej strednej hodnoty(anglicky conditional expectation constraints). Ak nastane porušenie obmedzení, teda $Z_i - T_i x > 0$, tak priemerná hodnota prekročenia obmedzenia bude menšia alebo rovná $d_i, i = 1, ..., r$.

Kapitola 3

Konvexita úloh pre spojité rozdelenia

V tomto oddieli sa budeme venovať konvexite úloh pravdepodobnostného programovania definovaných v kapitole 2. Za predpokladu, že poznáme rozdelenie všetkých náhodných veličín v danej úlohe, úloha sa niekedy môže riešiť prevedením na jej deterministický ekvivalent, ktorý je obyčajne typu (2.1)(kapitola 2,str.2) a na túto úlohu sa potom použije algoritmus pre riešenie úloh nelineárneho programovania. Väčšina algoritmov pre riešenie úlohy nelineárneho programovania (2.1) je založená na metóde vnútorného bodu t.j. na nasledujúcom princípe. Najprv nájdeme bod množiny prípustných riešení(na základe nutných podmienok pre optimálne riešenie). Potom sa postupne pohybujeme k ďalším prípustným bodom, v ktorých hodnota účelovej funkcie klesá. Pokračujeme v hľadaní vhodných smerov kým hodnota účelovej funkcie klesá. V prípade, že pre daný bod neexistuje smer, v ktorom by účelová funkcia klesala, vyhlasíme tento bod za kandidáta na optimum. Tieto metódy konvergujú iba k lokálnemu minimu. Je teda nutné celý algoritmus opakovať pre rôzne východzie body prípustných riešení.
V prípade, že účelová funkcia je konvexná na množine prípustných riešení, ktorá je tiež konvexná, potom lokálne minimum účelovej funkcie je zároveň aj globálnym minimom na množine prípustných riešení [1]. Konvexnosť úlohy sa teda ukazuje ako veľmi dôležitá vlastnosť úlohy nelineárneho programovania. Hlavnú rolu pre konvexitu množiny prípustných riešení úlohy (2.3)(kapitola 2, str. 3) bude hrať rozdelenie náhodného vektora Z. V ďalšom oddieli si pre spojité rozdelenia náhodného vektora Z ukážeme výsledok, ktorý zaručí konvexitu úlohy pre relatívne veľkú triedu úloh s pravdepodobnostnými obmedzeniami. Pre diskrétne rozdelenia ale tento výsledok platiť nebude a problém diskrétnych rozdelení si predstavíme osobitne v kapitole 4. Príklady algoritmov (vyžadujúcich konvexitu úlohy) na riešenie úloh s pravdepodobnostnými obmedzeniami sa dajú nájsť v ([2] str.287-298).
Zameriame sa aj na ostatné typy úloh definovaných v kapitole 2 (podtypy základnej úlohy (2.3)), na ktoré sa dajú aplikovať výsledky pre vyšetrovanie konvexity získané pre základnú úlohu a uvedieme aj ďalšie známe výsledky ohľadom kon-

vexity pre daný typ úlohy.

3.1 Postačujúce podmienky pre konvexitu množiny prípustných riešení pre spojité rozdelenia

V tejto sekcii predstavíme postačujúce podmienky pre konvexitu úlohy v základnom tvare (2.3)(kapitola 2, str. 3).Potom pre úlohu (2.5)(kapitola 2, str. 5) pre prípady náhodnej pravej strany a náhodnej matice. Osobitne pre Kataokov model (2.6)(kapitola 2, str. 5), kde je konvexita zaručená pre normálne rozdelenie a $p \geq \frac{1}{2}$. Nakoniec si predstavíme postačujúcu podmienku pre to, aby obmedzenia podmienenej strednej hodnoty v úlohe (2.8)(kapitola 2, str. 6) určovali konvexnú množinu.

3.1.1 Úloha v základnom tvare

Definície a vety 3.1-3.15 aj s dôkazmi sú prebraté z [3] (str. 94-112).

Definícia 3.1. *Povieme, že nezáporná funkcia $f(x)$ definovaná na konvexnej množine $\Omega \subseteq R^n$ je α-konkávna, kde $\alpha \in [-\infty, \infty]$, ak pre všetky $x, y \in \Omega$ a pre všetky $\lambda \in [0,1]$ platí nerovnosť:*

$$f(\lambda x + (1-\lambda)y) \geq m_\alpha(f(x), f(y), \lambda),$$

kde funkcia $m_\alpha : R_+ \times R_+ \times [0,1] \to R$ je definovaná ako

$$m_\alpha(a, b, \lambda) = 0 \quad ak \quad ab = 0,$$

a ak $a > 0, b > 0, 0 \leq \lambda \leq 1$, potom

$$m_\alpha(a,b,\lambda) = \begin{cases} a^\lambda b^{1-\lambda} & ak\ \alpha = 0 \\ \max(a,b) & ak\ \alpha = \infty \\ \min(a,b) & ak\ \alpha = -\infty \\ (\lambda a^\alpha + (1-\lambda) b^\alpha)^{\frac{1}{\alpha}} & inak. \end{cases}$$

Definícia 3.2. *Povieme, že nezáporná funkcia $f(x)$ definovaná na konvexnej množine $\Omega \subseteq R^n$ je log-konkávna, keď je 0-konkávna na tejto množine.*

Poznámka 3.3. *Ak f je kladná na Ω, f je log-konkávna práve vtedy, keď $\log f$ je konkávna na tejto množine.*

Definícia 3.4. *Povieme, že reálna funkcia $f : \Omega \to R$ definovaná na konvexnej množine $\Omega \subseteq R^n$ je kvázi-konkávna, keď $\forall x_1, x_2 \in \Omega, \forall \lambda \in [0,1]$ platí*

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \min(f(x_1, f(x_2)).$$

Poznámka 3.5. *Ak f je nezáporná, kvázi-konkavita je ekvivalentná s $-\infty$-konkavitou. Ak f je nezáporná, konkavita je ekvivalentná s 1-konkavitou. Ak f je konkávna na konvexnej množine $\Omega \subseteq R^n$, potom f je kvázi-konkávna na tejto množine.*

Poznámka 3.6. *Pre funkciu $f : R^2 \to R, f(x, y) = ax + by + k$, s parametrami $a \in R, b \in R, k \in R$ platí: $f(\lambda x_1+(1-\lambda)x_2, \lambda y_1+(1-\lambda)y_2) = \lambda(ax_1+by_1+k)+(1-\lambda)(ax_2 + by_2 + k), \forall(x_1, y_1) \in R^2, (x_2, y_2) \in R^2$. Z toho máme, že súčet lineárných funkcií je konkávna, kvázi-konkávna, konvexná funkcia na celom R^2.(Ľahko vieme tvrdenie rozšíriť na súčet n lineárnych funkcií.)*

Veta 3.7. *Nech funkcia $f : \Omega \to R$ definovaná na nejakej konvexnej množine $\Omega \subseteq R^n$ je kvázi-konkávna na tejto množine, potom pre každé $p \in R$ je množina $M := \{z \in \Omega : f(z) \geq p\}$ konvexná.*

Dôkaz
Vezmime $z_1 \in M, z_2 \in M, \lambda \in [0, 1]$. $M \subseteq \Omega$, takže $z_1 \in \Omega, z_2 \in \Omega$. Ω je konvexná, takže $\lambda z_1 + (1-\lambda)z_2 \in \Omega$. Z kvázi-konkavity f máme $f(\lambda z_1 + (1-\lambda)z_2) \geq \min(f(z_1), f(z_2)) \geq p$. Tzn. $\lambda z_1 + (1 - \lambda)z_2 \in M$.

QED.

.

Poznámka 3.8. *Platí aj opačná implikácia. Nech funkcia $f : \Omega \to R$ definovaná na nejakej konvexnej množine $\Omega \subseteq R^n$. Ak pre každé $p \in R$ je množina $M := \{z \in \Omega : f(z) \geq p\}$ konvexná potom f je kvázi-konkávna na tejto množine.*

Dôkaz
Nech $z_1, z_2 \in \Omega$. $p^* := min(f(z_1), f(z_2))$. Podľa predpokladu je $M_{p^*} = \{z \in \Omega : f(z) \geq p^*\}$ konvexná množina. Vezmime $\lambda \in [0, 1]$, potom $(\lambda z_1 + (1-\lambda)z_2) \in M_{p^*}$ a teda $f(\lambda z_1 + (1 - \lambda)z_2) \geq p^* = \min(f(z_1, f(z_2))$.

QED.

.

Veta 3.9. *Zobrazenie $\alpha \to m_\alpha(a, b, \lambda)$ je neklesajúce a spojité.*

Poznámka 3.10. *Veta 3.9 implikuje, že ak $f : \Omega \to R$, definovaná na nejakej konvexnej množine $\Omega \subseteq R^n$, je nezáporná α-konkávna funkcia na tejto množine, potom f je β - konkávna na tejto množine pre všetky $\beta \leq \alpha$.*

Definícia 3.11. *Povieme, že pravdepodobnostná miera P definovaná na Lebesgovsky meratetelných podmnožinách konvexnej množiny $\Omega \subset R^m$ je α-konkávna, ak pre všetky Borelovsky meratelné množiny $A, B \subset \Omega$ a pre všetky $\lambda \in [0, 1]$ platí nerovnosť*

$$P(\lambda A + (1 - \lambda)B) \geq m_\alpha(P(A), P(B), \lambda),$$

kde $\lambda A + (1-\lambda)B) = \{z \in R^m : z = \lambda x + (1-\lambda)y, x \in A, y \in B\}$
Povieme, že náhodný vektor Z *s hodnotami v* R^m *má* α*-konkávne rozdelenie, ak pravdepodobnostná miera* P_Z *indukovaná* Z *na* R^m *je* α*-konkávna.*

Predstavíme si teraz najobecnejšiu postačujúcu podmienku pre konvexitu množín s pravdepodobnostnými obmedzeniami. (Jedná sa o vetu 4.39 z [3].)

Veta 3.12. *Nech funkcie* $g_j : R^n \times R^m \to R, j \in J$ *sú kvázi-konkávne. Nech* Z *je* m*-rozmerný náhodný vektor, ktorý má* α*-konkávne rozdelenie, potom funkcia*

$$G(x) = P_Z\{g_j(x, Z) \geq 0, \forall j \in J\}$$

je α*-konkávna na množine*

$$D = \{x \in R^n : \exists z \in R^m \quad \textit{také, že} \quad g_j(x, z) \geq 0, j \in J\}.$$

Dôkaz([3] str. 107)
Nech $x_1, x_2 \in D, \lambda \in (0,1)$. Položme

$$\begin{aligned} A_1 &= \{z \in R^m : g_j(x_1, z) \geq 0, \forall j \in J\}, \\ A_2 &= \{z \in R^m : g_j(x_2, z) \geq 0, \forall j \in J\}, \\ B &= \lambda A_1 + (1-\lambda)A_2. \end{aligned}$$

P je α- konkávna miera na R^m, $A_1 \subseteq R^m$,$A_2 \subseteq R^m$:

$$P_Z(B) \geq m_\alpha(P_Z(A_1), P_Z(A_2), \lambda).$$

Vezmime $z \in B$. Potom z definície B máme:$\exists z_1 \in A_1, z_2 \in A_2$: $\lambda z_1 + (1-\lambda)z_2 = z$.
Z kvázikonkavity g_j platí:

$$g_j(\lambda x_1 + (1-\lambda)x_2, \lambda z_1 + (1-\lambda)z_2) \geq \min\{g_j(x_1, z_2), g_j(x_1, z_2)\} \geq 0, \forall j \in J.$$

Z tohto máme $\lambda x_1 + (1-\lambda)x_2 \in D$ a platí

$$B \subseteq \{z \in R^m : g_j(\lambda x_1 + (1-\lambda)x_2, z) \geq 0, \forall j \in J\}.$$

Spolu máme:

$$\begin{aligned} G(\lambda x_1 + (1-\lambda)x_2) &= P_Z(\{z \in R^m : g_j(\lambda x_1 + (1-\lambda)x_2, z) \geq 0, \forall j \in J\}) \\ &\geq P_Z(B) \geq m_\alpha(P_Z(A_1), P_Z(A_2), \lambda) = m_\alpha(G(x_1), G(x_2), \lambda). \end{aligned}$$

QED.

.

Poznámka 3.13. *Podľa dôkazu vety 3.12 množina* D *je konvexná.*

Poznámka 3.14. *Tento výsledok sa týka spojitých aj diskrétnych rozdelení pre náhodný vektor* Z*. Ďalší nástroj pre vyšetrovanie konvexity množiny* $\{x \in R^n | G(x) = P\{g_j(x, Z) \geq 0, j \in J\} \geq p\}$ *pre spojité rozdelenia vektora* Z *nám poskytne nasledujúce tvrdenie a jeho dôsledok.*

Veta 3.15. *Nech Ω je konvexná podmnožina R^m a nech $s > 0$ je dimenzia najmenšieho afinného podpriestoru L obsahujúceho Ω. Nech pravdepodobnostná miera P má hustotu vzhľadom k Lebesguovej miere na L. Potom pravdepodobnostná miera P na Ω je γ-konkávna s $\gamma \in [-\infty, \frac{1}{s}]$ práve vtedy, keď jej pravdepodobnostná hustota vzhľadom k Lebesgueovej miere na L je α - konkávna s*

$$\alpha = \begin{cases} \frac{\gamma}{1-s\gamma} & ak\ \gamma \in (-\infty, \frac{1}{s}), \\ -\frac{1}{s} & ak\ \gamma = -\infty, \\ +\infty & ak\ \gamma = \frac{1}{s}. \end{cases}$$

Poznámka 3.16. *Ak $\alpha = 0$ potom pravdepodobnostná miera je log-konkávna na Ω.*
Za podmienok vety 3.15 pravdepodobnostná miera P, ktorá má hustotu vzhľadom k Lebesguovej miere na R^m je kvázi-konkávna práve vtedy, keď jej pravdepodobnostná hustota je aspoň $-\frac{1}{m}$-konkávna.

Zaoberajme sa teraz postačujúcimi podmienkami pre konvexitu množiny prípustných riešení základnej úlohy pravdepodobnostného programovania (2.3)(kapitola 2,str. 3). Podľa vety 3.7 postačujúcou podmienkou k tomu, aby $M_i = \{x \in R^n : g_i(x) \geq p_i\}$ bola konvexná je kvázi-konkavita g_i. Ak teda $g_{r+1}, ..., g_m$ v úlohe (2.3) sú kvázi-konkávne, potom spolu vyčleňujú konvexnú množinu(prienik konvexných množín je konvexná množina). Nie je nutné, aby g_i boli kvázi-konkávne na celom priestore R^n. Ak ostatné obmedzenia pre množinu prípustných riešení vymedzujú konvexnú množinu M_{-i} potom je postačujúce, aby g_i boli kvázi-konkávne len na tejto množine.
Pravdepodobnostné obmedzenie vymedzuje množinu $M_0 = \{x \in R^n : h_0(x) \geq p\}$, $h_0(x) = P(g_1(x, Z) \geq 0, ..., g_r(x, Z) \geq 0)$. Ak $p > 0$, potom $M_0 = \{x \in D : h_0(x) \geq p\}$, kde D je ako z vety 3.12. Spojením viet 3.12 a 3.15 ku kvázi-konkavite h_0 na konvexnej množine D postačuje, aby $g_1, ..., g_r$ boli kvázi-konkávne na $R^n \times R^m$ a to, aby náhodný vektor Z mal spojité rozdelenie(vzhľadom k Lebesguovej miere na R^m) s hustotou $(-\frac{1}{s})$-konkávnou na R^m. Podľa poznámky 3.10 $(-\frac{1}{s})$ -konkavita hustoty je implikovaná log-konkavitou hustoty. Pre praktickú aplikáciu sa väčšinou overuje log-konkavita danej hustoty vektora Z.
Ak aj účelová funkcia $h(x)$ je konvexná, spolu máme konvexnú úlohu.

3.1.2 Prípady náhodnej pravej strany a náhodnej matice

Uvažujme teraz úlohu (2.5(kapitola 2,str. 5)), kde náhodný bude iba vektor Z. Ide o špeciálny prípad úlohy (2.3) a konvexita vyplynie z predchádzajúcej úvahy. Ak $F_Z(z) = P(Z \leq z)$ je distribučná funkcia náhodného vektora Z, ktorý by mal spojité rozdelenie s log-konkávnou hustotou, potom podľa viet 3.10,3.15 funkcia $F_Z(Tx)$ je log-konkávna(lineárne funkcie sú kvázi-konkávne) a teda množina $\{x \in R^n : P(Tx \geq Z) = F_Z(Tx) \geq p\}$ je konvexná pre ľubovoľné $p \in (0, 1)$ a úloha (2.5) je tak konvexná.
Môže nastať prípad, že hustota spojitého rozdelenia nebude log-konkávna a množina $\{x \in R^n : P(Tx \geq Z) = F_Z(Tx) \geq p\}$ bude predsa konvexná. Podľa vety 3.7 pre konvexnosť tejto množiny postačuje, aby distribučná funkcia $F_Z(z)$ bola kvázi-konkávna.

Poznámka 3.17. *Ak zložky náhodného vektora Z sú nezávislé a majú log-konkávne marginálne hustoty, potom aj funkcia združenej hustoty Z je log-konkávna. (Dôkaz by bol podobný ako u poznámky 4.1.) Ak ale zložky Z v takomto prípade nie sú nezávislé, log-konkavita združenej hustoty nie je zaručená. Prékopa ([2] str.281-282)uvádza postačujúce podmienky pre konvexitu úlohy (2.5) pre špeciálny prípad viacrozmerného združeného normálneho rozdelenia vektora Z, keď zložky náhodného vektora nie sú nezávislé.*

Pozrime sa teraz na ten prípad, keď sa náhoda v úlohe (2.5) objaví v matici. Pre prípad náhodnej matice so združeným normálnym rozdelením je známych niekoľko výsledkov týkajucich sa konvexity úlohy. ([2] str.285-286) Bez újmy na obecnosti vezmeme množinu určenú stochastickými obmedzeniami v tvare

$$\{x \in R^n : P(Tx \leq 0) \geq p\}, \tag{3.1}$$

kde T je matica $r \times n$. Nech T_{i*} značí i-tý riadok a T_{*j} j-tý stĺpec matice T. Nech μ_{i*} resp. μ_{*j} značia vektory strednýh hodnôt pre i-tý riadok resp. j-tý stĺpec matice T. Potom

Veta 3.18. *Nech prvky matice T majú združené normálne rozdelenie a pre korelačné matice medzi stĺpcami T platí:*

$$E\{(T_{*j} - \mu_{*j})(T_{*k} - \mu_{*k})^T\} = s_{jk}C,$$

kde C je konštantná kovariančná matica, $s_{jk} = s_{kj} \quad j, k = 1, ..., n$ sú konštanty, nech aj $p \geq \frac{1}{2}$, potom množina (3.1) je konvexná.

Veta 3.19. *Nech prvky matice T majú združené normálne rozdelenie a pre korelačné matice medzi riadkami T platí:*

$$E\{(T_{i*} - \mu_{i*})(T_{k*} - \mu_{k*})^T\} = s_{ik}C,$$

kde C je konštantná kovariančná matica, $s_{ik} = s_{ki} \quad i, k = 1, ..., r$ sú konštanty, nech aj $p \geq \frac{1}{2}$ potom množina (3.1) je konvexná.

V prípade, že množina prípustných riešení je daná ako

$$\{x \in R^n : P(Tx \leq Z) \geq p\}, \tag{3.2}$$

vieme ju upraviť na prienik konvexnej množiny a množiny s pravdepodobnostnými obmedzeniami ako v (3.1). Zavedieme novú maticu $(T, -Z)$ a nový rozhodovací vektor $(x^T, x_{n+1})^T$ a daná množina tvaru (3.2) je ekvivalentná množine $x \in R^n$ vektorov splňujúcich

$$P\Big((T, -Z)\begin{pmatrix} x \\ x_{n+1} \end{pmatrix} \leq 0\Big) \geq p$$
$$x_{n+1} = 1.$$

Obmedzenie v prvom riadku je v tom istom tvare ako v množine (3.1) a druhé obmedzenie určuje konvexnú množinu vektorov, takže ak aj prvé obmedzenie určí konvexnú množinu vektorov, potom množina daná týmito obmedzeniami je konvexná.(Prienik dvoch konvexných množín je konvexná množina.) Ak veta 3.18 alebo veta 3.19 platí pre obmedzenie v prvom riadku implikuje to zároveň aj konvexitu množiny danej obmedzeniami (3.2).
Posledný tu prezentovaný výsledok pre úlohu s náhodnou maticou sa týka prípadu log-normálneho rozdelenia prvkov náhodnej matice. Vyšetrujme znova množinu tvaru (3.2). Označme

$$G(x) = P(Tx \leq d) \geq p, \tag{3.3}$$

kde d je nenáhodný vektor.

Veta 3.20. *Nech náhodné členy $t_{i,k}$ $(i,k) \in I$ matice T sú kladné a združené pravdepodobnostné rozdelenie náhodného vektora $\alpha_{ik} = logt_{ik}, (i,k) \in I$ je log-konkávne. Nech náhodné členy matice T sa nachádzajú iba v prvých s stĺpcoch a všetky nenáhodné prvky matice v týchto stĺpcoch sú nezáponé. Potom funkcia*

$$G(e^{x_1}, ..., e^{x_s}, x_{s+1}, ..., x_n)$$

je log-konkávna.

Dôkaz

K dôkazu použijeme vetu 3.12. Predpokládajme $x_1, ..., x_s > 0$(A teda $x_1, ..., x_s$ sa budú dať reprezentovať ako $e^{z_1}, ..., e^{z_s}$). Položme

$$z_i = \log x_i, i = 1, ..., s$$
$$z_j = x_j, j = s+1, ..., n$$
$$z = (z_1, ..., z_n).$$

$P(t_{ij} > 0, (i,j) \in I) = 1$, položme

$$\alpha_{ij} = \log t_{ij}, (i,j) \in I.$$

Náhodný vektor α nech je vektor so zložkami $\alpha_{ij}, (i,j) \in I$.
Ďalej pre $k = 1, ..., r$ označme $I_k \subseteq \{1, ..., s\}$ množinu indexov takú, že $i \in I_k \Leftrightarrow (k,i) \in I(I_k$ bude označovať tie indexy i v k-tom riadku, že prvok matice t_{ki} je náhodný). Ďalej položme

$$g_k(z, \alpha) = d_k + \sum_{i \in I_k} (-e^{z_i + \alpha_{ki}}) + \sum_{j \in \{1,...,s\} \setminus I_k} (-t_{kj} e^{z_j}) + \sum_{j \in \{s+1,...,n\}} (-t_{kj} z_j).$$

Prvý člen d_k je konštanta. Je teda triviálne konkávnou funkciou.
Funkcia $y \to e^y$ je neklesajúca konvexná, lineárna funkcia $(z_i, \alpha_{ki}) \to z_i + \alpha_{ki}$ je konvexná podľa poznámky 3.6, takže funkcia

$$e^{z_i + \alpha_{ki}}$$

je konvexná podľa tvrdenia 2.5. Potom

$$-e^{z_i+\alpha_{ki}}$$

je konkávna funkcia, pre $i \in I_k$.
Keďže nenáhodné prvky matice v prvých s riadkoch sú nezáporné podľa predpokladu

$$-t_{kj}e^{z_j}$$

je konkávna funkcia pre $j \in \{1, ..., s\} \setminus I_k$.
Funkcia

$$(-t_{kj}z_j)$$

je znova konkávna funkcia, $j \in \{s+1, ..., n\}$, pretože lineárne funkcie sú konkávne. Spolu sa jedná o sumu konkávnych funkcií, takže g_k sú konkávne funkcie $k = 1, ..., r$. Podľa poznámky 3.5 g_k sú kvázi-konkávne $k = 1, ..., r$. Podľa predpokladov náhodný vektor α má log-konkávne rozdelenie, takže funkcia

$$H(z) = P(g_k(z, \alpha) \geq 0), k = 1, ..., r)$$

je log-konkávna na R^n podľa vety 3.12.
Nakoniec

$$\begin{gathered} G(e^{z_1}, ..., e^{z_s}, z_{s+1}, ..., z_n) = P(Tz \leq d) \\ = P\{t_{ij}, (i, j) \in I | t_{ij} > 0, (i, j) \in I, (T_{k*}z \leq d_k, k = 1, ..., r)\} \\ = P(g_k(z, \alpha) \geq 0), k = 1, ..., r) = H(z). \end{gathered}$$

QED.

3.1.3 Kataokov model

Množina prípustných riešení pre Kataokov model (2.6), keď vektor Z má n-rozmerné normálne rozdelenie, je konvexná pre $p \geq \frac{1}{2}$.([2] str.284-285)
Nech vektor $Z = (Z_1, ..., Z_n)$ má n-rozmerné normálne rozdelenie. Označme $\mu_i = \mathrm{E}(Z_i), i = 1, ..., n, \mu := (\mu_1, ...\mu_n)^T$ a $C := \mathrm{E}[(Z - \mu)(Z - \mu)^T]$.
Ak pre x máme $\mathrm{Var}(Z^T x) = x^T C x > 0$ potom platí

$$\mathrm{P}(Z^T x \geq d) = \mathrm{P}\Big(\frac{(Z-\mu)^T x}{\sqrt{x^T C x}} \geq \frac{d - \mu^T x}{\sqrt{x^T C x}}\Big) = 1 - \Phi\Big(\frac{d - \mu^T x}{\sqrt{x^T C x}}\Big)$$

Teda pravdepodobnostné obmedzenie v (2.6) je ekvivalentné

$$\mu^T x + \Phi^{-1}(1-p)\sqrt{x^T C x} \geq d.$$

Ak by $\mathrm{Var}(Z^T x) = 0$, potom toto obmedzenie by bolo taktiež ekvivalentné s pravdepodobnostným obmedzením. Problém (2.6) je teda ekvivalentný

$$\begin{gathered} \mathrm{Max}(\mu^T x + \Phi^{-1}(1-p)\sqrt{x^T C x}) \\ \text{za podmienok} \\ \textstyle\sum_{i=1}^n x_i = M, x \geq 0. \end{gathered} \tag{3.4}$$

Keďže C je pozitívne semidefinitná matica, funkcia $\sqrt{x^T C x}$ je konvexná. Keď $p \geq \frac{1}{2}$, potom $\Phi^{-1}(1-p) \leq 0$ a účelová funkcia v maximalizačnej optimalizačnej úlohe (3.4) je konkávna a teda úloha je konvexná.

3.1.4 Obmedzenia podmienených stredných hodnôt

V súvislosti s obmedzeniami podmienených stredných hodnôt platí([2] str.283):

Veta 3.21. *Ak Z je náhodná veličina, ktorá má spojité rozdelenie s log-konkávnou hustotou, potom*

$$g(z) = E(Z - z|Z - z > 0) \tag{3.5}$$

je klesajúca funkcia z.

V prípade, že prvky vektora Z v úlohe (2.8) majú každý marginálne rozdelenie, ktoré je spojité a log-konkávne, potom pre

$$g(Tx) = \mathrm{E}(Z - Tx|Z - Tx > 0) \leq d$$

existuje inverzná funkcia a obmedzenie vieme zapísať v ekvivalentnom tvare:

$$Tx \leq g^{-1}(d),$$

ktoré je lineárne a teda množina $x \in R^n$ určená týmto obmedzením je konvexná.

3.2 Kalkulus pre α-konkávne miery a funkcie

V tejto sekcii si zavedieme ďalšie tvrdenia, ktoré budú slúžiť ako kalkulus pre α-konkávne miery a funkcie. Vety 3.22-3.30 sú prebraté z ([3] str. 94-113), kde sú aj dokázané.

Veta 3.22. *Ak funkcia $f : R^n \to R_+$ je α-konkávna a funkcia $g : R^n \to R_+$ je β-konkávna, kde $\alpha, \beta \geq 1$, potom funkcia $h : R^n \to R_+$ definovaná ako $h(x) = f(x) + g(x)$ je γ-konkávna s $\gamma = min\{\alpha, \beta\}$.*

Nasledujúca veta je rozšírením vety (4.20 z[3]).

Veta 3.23. *Nech $g : R^n \to R^m$ je zobrazenie definované na konvexnej množine $\Omega \subseteq R^n$ a pre jeho zložky platí $g_i : R^n \to R, i = 1, ..., m$ sú konkávne funkcie na konvexnej množine Ω. Nech ďalej funkcia $F : R^m \to R$ je nezáporná neklesajúca α-konkávna na R^m, $\alpha \in [-\infty, \infty]$.Potom funkcia definovaná ako $F \circ g$ je α-konkávna na Ω.*

Dôkaz
Z konkávnosti g_i na Ω máme $\forall x_1, x_2 \in \Omega, \forall \lambda \in [0,1]$ platí:

$$g_i(\lambda x_1 + (1-\lambda)x_2) \geq \lambda g_i(x_1) + (1-\lambda) g_i(x_2), \forall i = 1, ..., m. \tag{3.6}$$

Z α-konkavity F máme $\forall y_1, y_2 \in R^m, \forall \lambda \in [0,1]$ platí:

$$F(\lambda y_1 + (1-\lambda)y_2) \geq m_\alpha(F(y_1), F(y_2), \lambda). \tag{3.7}$$

Vezmime ľubovoľné $\lambda \in [0,1], x_1, x_2 \in \Omega$. Platí:

$$F(g_1(\lambda x_1 + (1-\lambda)x_2), ..., g_m(\lambda x_1 + (1-\lambda)x_2)) \geq$$
$$F(\lambda g_1(x_1) + (1-\lambda)g_1(x_2), ..., \lambda g_m(x_1) + (1-\lambda)g_m(x_2)) \geq$$
$$m_\alpha(F(g(x_1)), F(g(x_2)), \lambda),$$

kde prvá nerovnosť je z toho, že F je neklesajúca na R^m (všetky súradnice klesnú podľa (3.6)). Druhá nerovnosť je z (3.7).

QED.

Veta 3.24. *Nech funkcia $f : R^m \times R^n \to R_+$ spĺňa, že pre všetky $y \in Y \subset R^n$ je funkcia $f(\cdot, y)$ α-konkávna na konvexnej množine $\Omega \subset R^m$, $\alpha \in [-\infty, \infty]$. Potom funkcia $\varphi(x) = \inf_{y \in Y} f(x,y)$ je α-konkávna na Ω.*

Veta 3.25. *Nech $\alpha_i > 0, i = 1, ..., m$ a $\sum\limits_{i=1}^{m} \alpha_i = 1$ potom funkcia $f : R_+^m \to R$, definovaná ako $f(x) = \prod\limits_{i=1}^{m} x_i^{\alpha_i}$ je konkávna.*

Veta 3.26. *Nech sú funkcie $f_i : R^n \to R_+, i = 1, ..., m$ α_i-konkávne a pre α_i platí $\sum\limits_{i=1}^{m} \alpha_i^{-1} > 0$, potom funkcia $g : R^n \to R_+$ definovaná ako $g(x) = \prod\limits_{i=1}^{m} f_i(x_i)$ je γ-konkávna s $\gamma = (\sum\limits_{i=1}^{m} \alpha_i^{-1})^{-1}$*

Veta 3.27. *Nech A je symetrická pozitívne definitná matica veľkosti $n \times n$, potom funkcia $A \to det(A)$ je $\frac{1}{n}$-konkávna.*

Veta 3.28. *Nech náhodný vektor Z indukuje α-konkávnu pravdepodobnostnú mieru na R^m, potom jeho združená distribučná funkcia F_Z je α-konkávna funkcia.*

Veta 3.29. *Nech náhodný vektor Z má navzájom nezávislé zložky s log-konkávnými marginálnymi rozdeleniami, potom Z ma log-konkávne rozdelenie.*

Veta 3.30. *Nech m-dimenzionálny náhodný vektor Z má α-konkávne rozdelenie, $\alpha \in [-\infty, \infty]$, a T je konštantná $n \times m$ matica. Potom n-dimenzionálny náhodný vektor $Y = TZ$ má α-konkávne rozdelenie.*

Veta 3.31. *Nech je dané zobrazenie $g : R^n \to R^m$ a nech každá zložka g_i je konkávna funkcia. Predokládajme, že náhodný vektor Z má navzájom nezávislé zložky a jednorozmerné marginálne distribučné funkcie $F_{Z_i}, i = 1, ...m$ sú α_i-konkávne. Ďalej, nech $\sum\limits_{i=1}^{k} \alpha_i^{-1} > 0$. Potom množina*

$$X_0 = \{x \in R^n : P\{g(x) \geq Z\} \geq p\}$$

je konvexná.

Veta 3.32 je prebratá z ([2] str.274).

Veta 3.32. *Ak pravdepodobnostná miera P je generovaná hustotou $f(z), z \in R^m$, ktorá má tú vlastnosť, že $f^{-\frac{1}{n}}(z), z \in R^m$ je konvexná funkcia, potom pravdepodobnostná miera P je kvázi-konkávna.*

3.3 Príklady spojitých rozdelení

Uvedieme príklady rozdelení, ktoré majú niektorú z požadovaných vlastností vyskytujúcich sa v postačujúcich podmienkach pre konvexitu úloh. Väčšina odvodení sa dá nájsť v ([2] str. 275-278).

1. **Nedegenerované normálne rozdelenie**
   Hustota nedegenerovaného normálneho rozdelenia je definovaná ako:

   $$f(x) = \frac{1}{\sqrt{|C|}(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)}, \quad x \in R^n,$$

   kde μ je vektor stredných hodnôt, C je kovariančná matica a $|C|$ značí determinant C. Predpokladáme, že C je pozitívne definitná matica. Potom aj C^{-1} je pozitívne definitná, a teda kvadratická forma $\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)$ je konvexná funkcia. $f(x) > 0, x \in R^n$ Potom

   $$\log f(x) = -\tfrac{1}{2}\log((2\pi)^n|C|) - \tfrac{1}{2}(x-\mu)^T C^{-1}(x-\mu)$$

   je konkávna funkcia a teda $f(x)$ je log-konkávna funkcia podľa poznámky 3.3.

2. **Dirichletovo rozdelenie**
   ([3] str. 98)Uvažujeme kladné čísla $\alpha_1, ..., \alpha_m$ a množinu

   $$S = \left\{ x \in R^m : \sum_{i=1}^{m} x_i \leq 1, x_i \geq 0, i = 1, ..., m \right\}.$$

   Hustota Dirichletovho rozdelenia s parametrami $\alpha_1, ..., \alpha_m$ je definovaná ako

   $$f(x) = \begin{cases} \dfrac{\Gamma(\alpha_1 + ... + \alpha_m)}{\Gamma(\alpha_1)...\Gamma(\alpha_m)} x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} ... x_m^{\alpha_m - 1} & \text{ak } x \in S, \\ 0 & \text{inak.} \end{cases}$$

   $\Gamma(\cdot)$ značí Gamma funkciu $\Gamma(z) = \int\limits_0^\infty t^{z-1}e^{-t}dt$. Predpokladajme, že $x_i > 0, i = 1, ..., m$, zlogaritmovaním hustoty dostaneme

   $$\log f(x) = \sum_{i=1}^{m}(\alpha_i - 1)\log x_i + \log\Gamma(\alpha_1 + ... + \alpha_m)) - \sum_{i=1}^{m}\log\Gamma(\alpha_i).$$

   Ak $\alpha_i \geq 1, i = 1, ..., m$, potom funkcia $\log f(\cdot)$ je konkávna funkcia. $f(\cdot)$ je teda log-konkávna pre $x \in S, x > 0$.

Nech teraz $x, y \in S$,$y > 0$, a predpokládajme, že existuje index $i \in \{1, ..., m\}$, že $x_i = 0$, vezmime $\lambda \in (0, 1)$. Máme

$$f(\lambda x + (1-\lambda)y) \geq 0 = f^{\lambda}(x)f^{1-\lambda}(y)$$

Pre $\lambda \in \{0, 1\}$ nerovnosť platí triviálne.
Takisto vieme nerovnosť dokázať aj pre dvojicu $x, y \in S$: existujú indexy i, j, že $x_i = 0, y_j = 0$.
$f(x)$ je teda log-konkávna na S pre $\alpha_i \geq 1, i = 1, ..., m$.

3. **Jednorozmerné gamma rozdelenie**
Uvažujme hustotu:

$$f(z) = \begin{cases} \dfrac{\lambda^{\vartheta} z^{\vartheta-1} e^{-\lambda z}}{\Gamma(\vartheta)} & \text{pre} \quad z > 0 \\ 0 & \text{inak} \end{cases}$$

kde $\lambda > 0$ a $\vartheta > 0$ sú konštanty. Ak $\lambda = 1$ povieme, že rozdelenie je štandartdné gamma rozdelenie. Ak náhodná premenná Y má gamma rozdelenie, ϑY má štandardné gamma rozdelenie. Pre $z > 0$ máme:

$$\log f(z) = -.\log(\Gamma(\vartheta)) + \log(\lambda^{\vartheta}) + (\vartheta - 1)\log(z) - \lambda z$$

Funkcia je $\log f(\cdot)$ konkávna pre $\vartheta \geq 1$, teda v tom istom prípade je $f(\cdot)$ aj log-konkávna pre $z > 0$.

4. **Viacrozmerné gamma rozdelenie**
Toto rozdelenie môžeme definovať určitou transformáciou m nezávislých náhodných veličín $Z_1, ..., Z_m$, ktoré majú štandartné gamma rozdelenie. Nech je daná matica A $(n \times m)$ s prvkami 0 alebo 1. Položme $Z = Z_1, ..., Z_m$ a definujme

$$Y = AZ$$

Potom náhodný vektor Y má štandardné viacrozmerné gamma rozdelenie. Rozdelenie vektora Z je log-konkávne podľa vety 3.29. A tak n-rozmerné štandardné gamma rozdelenie je log-konkávne podľa vety 3.30.

5. **Wishartovo rozdelenie**
Nech $M > s, M \in N, s \in N$, potom Wishartovo rozdelenie je spojité rozdelenie na priestore symetrických štvorcových matíc s pravdepodobnostnou hustotou:

$$f(A) = \begin{cases} \dfrac{\det(A)^{\frac{M-s-2}{2}} \exp(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}A))}{2^{\frac{M-1}{2}s} \pi^{\frac{s(s-1)}{4}} \det(\Sigma)^{\frac{M-1}{2}} \prod_{i=1}^{s} \Gamma(\frac{M-i}{2})} & \text{pre A pozitívne definitnú,} \\ 0 & \text{inak,} \end{cases}$$

kde A je matica $(s \times s)$, Σ je konštantná matica $(s \times s)$.
Nech A_1, A_2 sú pozitívne definitné matice $(s \times s)$, $\lambda \in [0,1]$. Z vlastností stopy matice máme

$$\begin{aligned} \mathrm{tr}\Sigma^{-1}(\lambda A_1 + (1-\lambda)A_2) &= \mathrm{tr}[\lambda\Sigma^{-1}A_1 + (1-\lambda)\Sigma^{-1}A_2] \\ &= \lambda\mathrm{tr}\Sigma^{-1}A_1 + (1-\lambda)\mathrm{tr}\Sigma^{-1}A_2. \end{aligned}$$

$A \to \mathrm{tr}\Sigma^{-1}A$ je teda konvexná funkcia.
Zlogaritmovaním hustoty $f(A)$ dostávame

$$\log f(A) = C + \tfrac{1}{2}(M - s - 2)\log\det A - \tfrac{1}{2}\mathrm{tr}\Sigma^{-1}A,$$

kde $C \in R$ nie je funkciou A. Podľa vety 3.27 $A \to \det(A)$ je $\frac{1}{s}$-konkávna funkcia a podľa poznámky 3.10 je aj log-konkávna. Teda funkcia $A \to \log\det(A)$ definovaná na množine pozitívne definitných štvorcových matíc je konkávna funkcia a $A \to -\mathrm{tr}\Sigma^{-1}A$ je taktiež konkávna. Z toho ihneď vidieť, že ak $M \geq s+2$, potom f je log-konkávna funkcia.

6. **Log-normálne rozdelenie**
Hustota jednorozmerného log-normálneho rozdelenia s parametrami μ a σ je definovaná ako:

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x}\exp\Big(-\frac{(\ln x - \mu)^2}{2\sigma^2}\Big) & \text{pre} \quad x > 0 \\ 0 & \text{inak} \end{cases}$$

Ukážeme, že združená distribučná funkcia je log-konkávna.
m-dimenzionálny náhodný vektor Z má log-normálne rozdelenie ak vektor $Y = (\log Z_1, ..., \log Z_m)^T$ má viacrozmerné normálne rozdelenie.(Normálne rozdelenie je log-konkávne.) Distribučná funkcia Z v bode $z \in R^m, z > 0$ je:

$$F_Z(z) = \mathrm{P}(Z_1 \leq z_1, ..., Z_m \leq z_m) = P(z_1 - e^{Y_1} \geq 0, ..., z_m - e^{Y_m} \geq 0)$$

Predpoklady vety 3.12 sú splnené pre pravú stranu, kde $g_i(x, Y) = z_i - e^{Y_i}$ a teda F_Z je log-konkávna funkcia.
Log-konkávnosť distribučnéj funkcie rozdelenia neimplikuje log-konkavitu hustoty ani log-konkavitu rozdelenia, ale ako sme si ukázali(sekcia 3.1.2) log-konkavita distribučnej funkcie náhodného vektora Z je postačujúca podmienka pre zaručenie konvexity úlohy 2.5 podľa vety 3.7.

7. **Cauchyho rozdelenie**
Je definované ako združené rozdelenie náhodných veličín

$$\zeta_i = \frac{\sqrt{\nu}Z_i}{\eta}, \quad i = 1, ..., n,$$

kde $(Z_1, ..., Z_n)$ má štandardné normálne rozdelenie, $(Z_1, ..., Z_n)$ je nezávislé s náhodnou veličinou η, ktorá má χ^2-rozdelenie s μ stupňami voľnosti. Hustota tohoto rozdelenia je:

$$f(x) = \frac{\Gamma(\dfrac{1}{2}(\mu+n))}{(\pi\mu)^{\frac{n}{2}}\Gamma(\dfrac{1}{2}\mu)|R|^{\frac{1}{2}}}(1+\tfrac{1}{\mu}x^T R^{-1}x)^{-\frac{1}{2}(\mu+n)}$$

pre $x \in R^n$.
n-rozmerná Cauchyho hustota má tú vlastnosť, že $f^{-\frac{1}{n}}$ je konvexná funkcia v R^n Podľa vety 3.32 je toto rozdelenie kvázi-konkávne. Cauchyho rozdelenie je teda také, aké požadujeme v predpokladoch vety 3.12 s $\alpha = -\infty$.

8. **Paretovo rozdelenie**
Hustota tohoto rozdelenia je definovaná ako:

$$f(x) = a(a+1)...(a+n-1)\Big(\prod_{j=1}^{n}\Theta_j\Big)^{-1}\Big(\sum_{j=1}^{n}\Theta_j^{-1}x_j - n + 1\Big)^{-(a+n)}$$

pre $x_i > \Theta_i, i = 1, ..., n,$ $f(x) = 0$ inak; $\Theta_1, ..., \Theta_n$ sú kladné konštanty. Dá sa ľahko ukázať, že $f^{-\frac{1}{n}}$ je konvexná funkcia v R^n. Podľa vety 3.32 je teda toto rozdelenie kvázi-konkávne a je to teda rozdelenie z predpokladov vety 3.12 s $\alpha = -\infty$.

Uvedieme si ešte jeden príklad rozdelenia s log-konkávnou hustotou bez dôkazu. Naznak dôkazu pre log-konkavitu tejto funkcie sa dá nájsť v [2] (str.276)

9. **Beta rozdelenie**
Má hustotu definovanú na množine pozitívne definitných štvorcových matíc:

$$f(X) = \frac{c(n_1,p)c(n_2,p)}{c(n_1+n_2,p)}|X|^{\frac{1}{2}(n_1-p-1)}|I-X|^{\frac{1}{2}(n_2-p-1)}$$

kde $X, I-X$ sú pozitívne definitné matice $p \times p$ a

$$\frac{1}{c(k,p)} = 2^{\frac{pk}{2}}\pi^{\frac{p(p-1)}{2}}\prod_{i=1}^{p}\Gamma(\frac{k-i+1}{2}).$$

3.4 Riešené príklady

Príklad 3.33. *Nech Z, W sú nezávislé náhodné veličiny, Z so štandardným normálnym rozdelením, W s χ^2 rozdelením o jednom stupni voľnosti. A buď pozitívne definitná matica $n \times n$. Je množina $M = \{x \in R^n : P(Z - \sqrt{(x^T Ax)W}) \geq 0) \geq p\}$, $p \in (0,1)$ konvexná?*

Riešenie

$$\begin{aligned} M &= \{x \in R^n : P(Z \geq \sqrt{(x^T Ax)}\sqrt{W})) \geq p\} \\ &= \{x \in R^n : P(\tfrac{Z}{\sqrt{W}} \geq \sqrt{(x^T Ax)})) \geq p\} \end{aligned}$$

W, Z sú nezávislé náhodné veličiny, takže veličina $U = \frac{Z}{\sqrt{W}}$ má t rozdelenie o 1 stupni voľnosti.

$$\begin{aligned} M &= \{x \in R^n : 1 - F_U(\sqrt{(x^T Ax)}) \geq p\} \\ &= \{x \in R^n : t_1^{-1}(1-p) \geq \sqrt{(x^T Ax)}\} \\ &= \{x \in R^n : [t_1^{-1}(1-p)]^2 \geq x^T Ax\} \\ &= \{x \in R^n : -[t_1^{-1}(1-p)]^2 \leq -x^T Ax\}, \end{aligned}$$

kde t_1^{-1} je kvantilová funkcia t-rozdelenia s 1 stupňom voľnosti. $x^T Ax$, kde A je pozitívne definitná, je konvexná funkcia(Hessián tejto funkcie je pozitívne definitný), potom $-x^T Ax$ je konkávna funkcia. Konkávnosť implikuje kvázi-konkavitu $-x^T Ax$ podľa poznámky 3.5. M je konvexná podľa vety 3.7.

Príklad 3.34. *Predpokládajme, že P je α- konkávne pravdepodobnostné rozdelenie a $A \subset R^n$ je konvexná množina. Dokážeme, že funkcia $f(x) = P(A + x)$ je α-konkávna.(Úloha je vyriešená v))*

Riešenie:
P je α-konkávna miera na R^n. Pre všetky Borelovsky meratelné množiny $B_1, B_2 \subset R^n$,$\forall\lambda \in [0, 1]$ platí:

$$P(\lambda B_1 + (1-\lambda)B_2) \geq m_\alpha(P(B_1), P(B_2), \lambda)$$

Vezmime teraz $\lambda \in [0, 1], x \in A$. Hľadáme $y_1, y_2 \in A : x = \lambda y_1 + (1-\lambda)y_2$, voľme $y_1 = x, y_2 = x$, zrejme $x = \lambda x + (1-\lambda)x$. Máme $A \subseteq (\lambda A + (1-\lambda)A)$.
Nech $\lambda \in [0, 1], x \in (\lambda A + (1-\lambda)A)$. Potom existujú $y_1, y_2 \in A : x = \lambda y_1 + (1-\lambda)y_2$. Potom $x \in A$ z konvexity A. $(\lambda A+(1-\lambda)A) \subseteq A$. $\lambda \in [0, 1]$ bolo ľubovoľné, takže máme:

$$A = \lambda A + (1-\lambda)A, \forall\lambda \in [0, 1].$$

Nech teraz $\lambda \in [0, 1], x, y \in R^n$. Platí $A = \lambda A + (1-\lambda)A$, takže môžeme písať:

$$\begin{aligned} f(\lambda x + (1-\lambda)y) = P(A + \lambda x(1-\lambda)y) = P(\lambda(A + x) + (1-\lambda)(A + y)) \\ \geq m_\alpha((A + x), P(A + y), \lambda) = m_\alpha(f(x), f(x), \lambda). \end{aligned}$$

QED.

.

Príklad 3.35. *Dokážeme,že ak $\theta : R \to R$ je log-konkávna pravdepodobnostná hustota, potom funkcia $G(x) = 1 - \int\limits_{t \leq x} \theta(t)dt$ je log-konkávna.*

Riešenie:
Podľa vety 3.15 pravdepodobnostná miera určená hustotou θ je log-konkávna. Ďalej máme $G(x) = P(\theta - x \geq 0)$. Podľa poznámky (3.6) $f(x, \theta) = (\theta - x)$ je kvázi-konkávna. Podľa vety 3.12 dostávame log-konkavitu $F(x)$.

QED.

Ďalšie tvrdenie by sa dalo dokázať obdobne, my ho ale dokážeme priamo. Použijeme k tomu obdobu Brunn-Minkovského nerovnosti z ([5] str.3).Prékopa pomocou tejto nerovnosti naznačil dôkaz tvrdenia (riešenie príkladu) v ([5] str.5).

Veta 3.36. *Nech $f_1, ..., f_k$ sú nezáporné a Borelovsky meratelné funkcie definované na R^n a nech*

$$r(t) = \sup\{f_1(x_1)...f_k(x_k) : \lambda_1 x_1 + ... + \lambda x_k = t, x_i \in R^n, i = 1, ..., k\}, t \in R^n,$$

kde $\lambda_1, ..., \lambda_k$ sú kladné konštatnty také, že $\lambda_1 + ... + \lambda_k = 1$. Potom funkcia $r(t), t \in R^n$ je tiež borelovský meratelná a platí nerovnosť:

$$\int_{R^n} r(t)dt \geq (\int_{R^n} f^{\frac{1}{\lambda_1}}(x_1))^{\lambda_1}...(\int_{R^n} f^{\frac{1}{\lambda_k}}(x_k))^{\lambda_k}.$$

Príklad 3.37. *Dokážeme,že ak $\theta : R \to R$ je log-konkávna pravdepodobnostná hustota, potom funkcia $F(x) = \int_{t \leq x} \theta(t)dt$ je log-konkávna.*

Riešenie:
Dokážeme, že pravdepodobnostná miera generovaná hustotou $\theta(t)$ je log-konkávna. Nech $A, B \subset R$ sú Borelovsky meratelné, $\lambda \in (0, 1)$. Môžeme definovať:

$$f_1(t) = \theta^{\lambda}(t) \text{ pre } t \in A, f_1(t) = 0 \text{ inak},$$
$$f_2(t) = \theta^{1-\lambda}(t) \text{ pre } t \in B, f_2(t) = 0 \text{ inak},$$
$$r(t) = \sup\{f_1(x_1)f_2(x_2) : \lambda x_1 + (1-\lambda)x_2 = t, x_1 \in R, x_2 \in R\}, t \in R.$$

Z log-konkavity $\theta(t)$ na R platí:

$$\theta(\lambda x_1 + (1-\lambda)x_2) \geq \theta^{\lambda}(x_1)\theta^{1-\lambda}(x_2), x_1 \in R, x_2 \in R.$$

$\forall(t \in R), \exists(x_1 \in R, x_2 \in R) : \lambda x_1 + (1-\lambda)x_2 = t$, takže máme nerovnosť:

$$\theta(t) \geq \sup\{\theta^{\lambda}(x_1)\theta^{1-\lambda}(x_2) : \lambda x_1 + (1-\lambda)x_2 = t, x_1 \in R, x_2 \in R\}, \forall t \in R.$$

Nech teraz $t \notin (\lambda A + (1-\lambda)B)$. Ak $\lambda x_1 + (1-\lambda)x_2 = t, x_1 \in R, x_2 \in R$ potom platí: ($x_1 \notin A$ a $f_1(x_1) = 0$) alebo ($x_2 \notin B$ a $f_2(x_2) = 0$). Takže $r(t) = 0$ pre $t \notin (\lambda A + (1-\lambda)B)$ a máme rovnosť

$$\int\limits_{\lambda A + (1-\lambda)B} r(t)dt = \int\limits_{R^n} r(t)dt.$$

Z tohto a podľa vety 3.36 dostávame:

$$\begin{aligned} P(\lambda A + (1-\lambda)B) &= \int\limits_{\lambda A + (1-\lambda)B} \theta(t)dt \geq \int\limits_{\lambda A + (1-\lambda)B} r(t)dt = \\ &= \int\limits_{R^n} r(t)dt \geq \Big(\int\limits_{R^n} f_1^{\frac{1}{\lambda}}(t)dt\Big)^{\lambda}\Big(\int\limits_{R^n} f_2^{\frac{1}{1-\lambda}}(t)dt\Big)^{1-\lambda} = \\ &= \Big(\int\limits_{R^n} \theta(t)dt\Big)^{\lambda}\Big(\int\limits_{R^n} \theta(t)dt\Big)^{1-\lambda} = P^{\lambda}(A)P^{1-\lambda}(B). \end{aligned}$$

Pre $\lambda \in \{0, 1\}$ nerovnosť platí triviálne. Dokázali sme, že pravdepodobnostná miera P je log-konkávna na R. Podľa vety 3.28, distribučná funkcia $F(x) = \int_{t \leq x} \theta(t)dt$ je tiež log-konkávna.

QED.

Príklad 3.38. *Ak Y je m-rozmerný náhodný vektor s log-normálnym rozdelením, a $g : R^n \to R^m$ je také zobrazenie, že každá komponenta $g_i, i = 1, .., m$ je konkávna funkcia, dokážeme,že množina*

$$C = \{x \in R^n : P(g(x) \geq Y) \geq 0,9\}$$

je konvexná.

Riešenie:

$$C = \{x \in R^n : F_Y(g(x)) \geq 0,9\}.$$

Distribučná funkcia $F_Y(y)$ viacrozmerného log-normálneho rozdelenia(dôkaz uvedený v sekcii s príkladmi spojitých rozdelení) je log-konkávna nezáporná a neklesajúca. Podľa vety 3.23 funkcia $F_Y(g(x))$ je log-konkávna a C je konvexná množina podľa vety 3.7.

Príklad 3.39. *Nech náhodná veličina Z má rovnomerné rozdelenie na intervale $[0, 1]$ a $e = (1, ..., 1)^T$. Dokážeme,že nasledujúca množina je konvexná:*

$$C = \{x \in R^n : P(\exp(x^T y) \geq e^T y Z, \forall y \in R^n : ||y|| \leq 1) \geq 0,95\}.$$

Riešenie:
Pre $y = 0$ je nerovnosť v pravdepodobnostnom obmedzení splnená s pravdepodobnosťou 1, preto:

$$C = \Big\{x \in R^n : P\Big((\frac{\exp(x^T y)}{e^T y} \geq Z, \forall y \in R^n : e^T y > 0, ||y|| \leq 1) \\ \wedge (\frac{\exp(x^T y)}{e^T y} \leq Z, \forall y \in R^n : e^T y < 0, ||y|| \leq 1) \Big) \geq 0,95 \Big\}.$$

Druhá podmienka je splnená s pravdepodobnosťou 1. Preto môžeme písať:

$$C = \Big\{x \in R^n : P\Big(\frac{\exp(x^T y)}{e^T y} \geq Z, \forall y \in R^n : e^T y > 0, ||y|| \leq 1 \Big) \geq 0,95 \Big\}.$$

0značme

$$f(x, y) = \frac{\exp(x^T y)}{e^T y}, x, y \in R^n,$$
$$g(x) = \inf \Big\{ f(x, y), y \in R^n : e^T y > 0, ||y|| \leq 1 \Big\}.$$

Potom

$$C = \{x \in R^n : P(g(x) \geq Z) \geq 0,95\} = \\ \{x \in R^n : g(x) \geq 0,95\}.$$

Zafixujeme $y^* \in R^n$. $\log(f(x, y^*)) = x^T y^* - \log(e^T y^*)$. Podľa poznámky 3.6 je toto konkávna funkcia x na R^n. Funkcia $f(\cdot, y^*) : R^n \to R_+$ je teda log-konkávna pre všetky $y^* \in R^n$ a to teda platí aj pre všetky $y^* \in \{y \in R^n : e^T y > 0, ||y|| \leq 1\}$. Podľa vety 3.24 $g(x)$ je log-konkávna na R^n(a teda aj kvázi-konkávna podľa poznámky 3.10) a C je konvexná množina podľa vety 3.7.

QED.

Kapitola 4

Diskrétne rozdelenia

Diskrétne rozdelenia majú využitie jednak pre náhodné veličiny, ktoré môžu nadobúdať iba celočíselné hodnoty, napríklad počet ľudí, množstvo nakupovanej komodity, jednak na aproximáciu spojitých veličín celými číslami, napríklad množstvo spotrebovanej energie v kWh.

4.1 Postačujúce podmienky pre konvexitu pre diskrétne rozdelenia

V tejto kapitole sa obmedzíme na úlohu tvaru, ktorá má náhodnú iba pravú stranu:

$$\begin{gathered} \min c(x) \\ \text{za podmienok} \\ \mathrm{P}(g(x) \geq Z) \geq p \\ x \in X, \end{gathered} \tag{4.1}$$

kde predpokladáme, že $X \subset R^n$ je uzavretá konvexná množina, $c : R^n \to R$ je konvexná funkcia, Z je m-rozmerný náhodný vektor a $g_i : R^n \to R$ sú konkávne funkcie $i = 1, ..., m$.
Pravdepodobnostné obmedzenie vieme vyjadriť ako obmedzenie distribučnej funkcie:

$$\mathrm{P}(g(x) \geq Z) \geq p \Leftrightarrow F_Z(g(x)) \geq p.$$

Pre diskrétne rozdelenia existuje analógia definície log-konkávnosti pre postupnosti.
([3] str.106)

Definícia 4.1. *Povieme, že postupnosť $\{p_k | k \in Z\}$ je log-konkávna, keď platí*

$$p_k^2 \geq p_{k-1} p_{k+1}, \forall k \in Z.$$

Veta 4.2. *Nech náhodná veličina Z má celočíselné rozdelenie $P(Z = k) = p_k, k \in Z$, a $\{p_k, k \in Z\}$ tvorí log-konkávnu postupnosť. Potom distribučná funkcia F_Z je log-konkávna funkcia.*

Veta 4.2 ([3], str. 106) nám pomáha zodpovedať otázku ohľadom konvexity úlohy (4.1) pre jednorozmerný vektor Z. Podľa vety 3.23 a vety 3.7 , ak $g : R \to R$

je konkávna funkcia a Z je náhodná veličina s rozdelením spĺňajúcim predpoklady vety 4.2 potom množina $\{x \in R^n : F_Z(g(x)) \geq p\}$ je konvexná.

Poznámka 4.3. *Tvrdenie vieme rozšíriť aj pre prípad, keď každá zo zložiek náhodného vektora je log-konkávna a jeho zložky sú navzájom nezávislé. Keď sú zložky $Z_i, i = 1, ...m$ náhodného vektora Z navzájom nezávislé, každá s marginálnym diskrétnym rozdelením, ktoré tvorí log-konkávnu postupnosť ako je definovaná v definícii 4.1 a g_i sú konkávne funkcie $i = 1, ..., m$, jednoduchou úvahou vieme dospieť k výsledku, že aj v takom prípade je množina $\{x \in R^n : F_Z(g(x)) \geq p\}$ konvexná.*

Dôkaz
Nezávislosť zložiek vektora Z implikuje

$$F_Z(z_1, ..., z_m) = \prod_{i=1}^{m} F_{Z_i}(z_i).$$

Vezmime $z, y \in R^m$, $\lambda \in [0, 1]$, potom platí

$$F_Z(\lambda z + (1 - \lambda)y) = \prod_{i=1}^{m} F_{Z_i}(\lambda z_i + (1 - \lambda)y_i)$$

$$\geq \prod_{i=1}^{m} F_{Z_i}^{\lambda}(z_i)F_{Z_i}^{1-\lambda}(y_i) = (\prod_{i=1}^{m} F_{Z_i}(z_i))^{\lambda}(\prod_{i=1}^{m} F_{Z_i}(y_i))^{1-\lambda} = F_Z^{\lambda}(z)F_Z^{1-\lambda}(y),$$

kde nerovnosť plynie z nezápornosti a log-konkavity $F_{Z_i}, i = 1, ..., m$. F_Z je neklesajúca log-konkávna funkcia na R^m, $g_i(x), i = 1, ..., m$ sú konkávne na R^n spojením viet 3.23 a 3.7 dostávame konvexnosť množiny $\{x \in R^n : F_Z(g(x)) \geq p\}$.

QED.

Uveďme si príklady jednorozmerných diskrétnych rozdelení, spĺňajúcich vlastnosť log-konkavity pre postupnosti.

1. **Binomické rozdelenie**
   je definované ako

$$p_k = \mathrm{P}(K = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad n \geq k \geq 1, n, k \in N,$$
$$p_k = 0, \text{inak.}$$

   platí známa nerovnosť

$$\binom{n}{k}^2 \geq \binom{n}{k-1}\binom{n}{k+1} \quad \forall n \geq k \geq 1, \tag{4.2}$$

   Z toho ihneď máme $p_k^2 \geq p_{k-1}p_{k+1}, k \in Z$.

2. **Geometrické rozdelenie**
je definované ako

$$p_k = \mathrm{P}(K = k) = p^k(1-p)^k, \quad k \geq 1, k \in Z,$$
$$p_k = 0, \quad \text{inak.}$$

Zrejme

$$p_k^2 = p_{k-1}p_{k+1}, \forall k \geq 2, k \in Z,$$
$$p_k^2 \geq p_{k-1}p_{k+1} = 0, \forall k < 2, k \in Z,$$

Spolu máme

$$p_k^2 \geq p_{k-1}p_{k+1}, \forall k \in Z,$$

čo implikuje log-konkavitu postupnosti.

3. **Hypergeometrické rozdelenie**
Pre

$$N \in \{0, 1, 2, ...\},$$
$$K \in \{0, 1, 2, ..., N\},$$
$$n \in \{0, 1, 2, ..., N\},$$

je definované ako

$$p_k = \mathrm{P}(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}},$$

pre $k \in N : \max(0, n + K - N) \leq k \leq \min(K, n)$, $p_k = 0$, inak.
Overíme log-konkavitu postupnosti. Podľa (4.2) platí

$$\binom{K}{k}^2 \geq \binom{K}{k+1}\binom{K}{k-1},$$
$$\binom{N-K}{n-k}^2 \geq \binom{N-K}{n-k+1}\binom{N-K}{n-k-1},$$

pre $k \in N : \max(0, n + K - N) + 1 \leq k \leq \min(K, n) - 1$. Spolu máme

$$\binom{K}{k}^2\binom{N-K}{n-k}^2 \geq \binom{K}{k+1}\binom{N-K}{n-k-1}\binom{K}{k-1}\binom{N-K}{n-k+1}$$

a teda

$$p_k^2 \geq p_{k-1}p_{k+1},$$

pre $k \in N : \max(0, n+K-N)+1 \leq k \leq \min(K, n)-1$. Pre ostatné $k \in Z$ platí nerovnosť triviálne a spolu máme

$$p_k^2 \geq p_{k-1}p_{k+1}, k \in Z.$$

4. **Poissonovo rozdelenie**
je pre $\lambda > 0$ definované ako

$$p_k = \mathrm{P}(K = k) = \frac{\lambda^k e{-}\lambda}{k!}, \quad k = 0, 1...., \\ p_k = 0, \quad \text{inak.}$$

Overíme log-konkavitu postupnosti. Zrejme platí

$$(k+1)(k-1)! \geq k(k-1)!, \\ \frac{1}{(k+1)(k-1)!} \leq \frac{1}{k!}, \\ \frac{\lambda^{k+1}}{(k+1)k!}e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda} \leq \frac{\lambda^{2k}}{k!k!}e^{-\lambda}, \\ p_k^2 \geq p_{k-1}p_{k+1},$$

pre $k = 1, 2, ...$ Pre $k \leq 0$ nerovnosť platí triviálne.

4.2 Nekonvexnosť úlohy

V predchádzajúcej kapitole sme si predstavili postačujúce podmienky pre konvexitu úloh s pravdepodobnostnými obmedzeniami v prípade, že Z má spojité rozdelenie. V prípade, že Z má diskrétne rozdelenie, je situácia zložitejšia a konvexita nie je zaručená ani za pomerne silných podmienok. Pre diskrétne rozdelenia sme si vyčlenili úlohy, ktoré sú konvexné, no na rozdiel od prípadov so spojitými rozdeleniami, nepokryjú veľkú časť praktických úloh. (Napríklad nemôžeme čakať, že zložky náhodného vektora budú nezávislé.) Riešenie sa ponúka pomocou p-level eficientných bodov (anglicky p-level efficient points, PLEP), pomocou ktorých dokážeme riešiť aj nekonvexné úlohy rozdelením úlohy na niekoľko konvexných úloh.

Pre ilustráciu si predstavíme nasledujúci príklad:

Príklad 4.4.

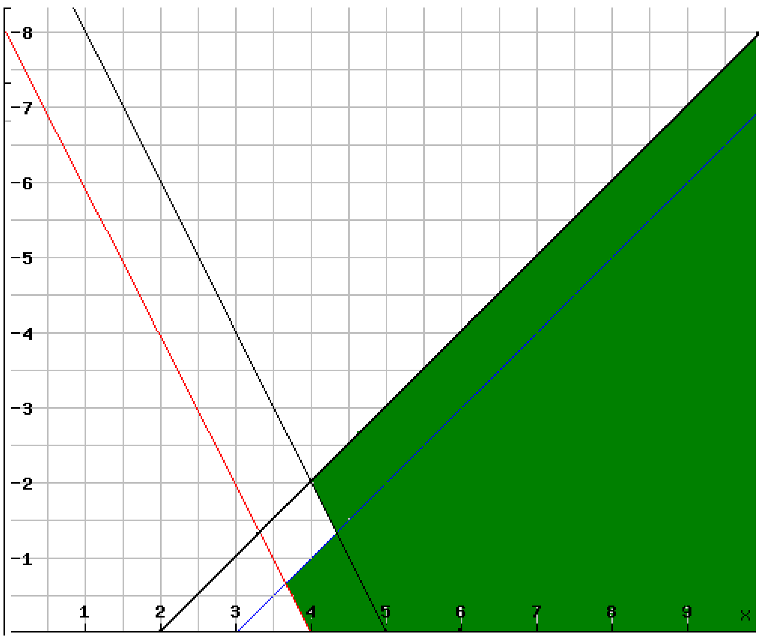
$$\begin{gathered} \mathit{min} \quad x_1 \\ \mathit{za\ podmienok} \\ x_1 - x_2 \geq z_1 \\ x_1 + \tfrac{1}{2}x_2 - 2 \geq z_2 \\ x_1 \geq 0, x_2 \geq 0 \end{gathered} \tag{4.3}$$

kde splnenie zároveň oboch obmedzení požadujme s pravdepodobnosťou aspoň $p = 0,55$. *Rozdelenie vektora Z je dané tabuľkou (tabuľka 4.1).*

Na obrázku (obr. 4.1) je vyznačená množina prípustných riešení úlohy, ktorá ani pre túto jednoduchú úlohu zrejme nie je konvexná. S nekonvexnosťou množiny prípustných riešení sa pri praktických úlohách s diskrétnymi obmedzeniami stretneme častejšie a na konvexitu úlohy sa teda nemôžeme veľmi spoliehať. Pri riešení úlohy sa treba vysporiadať s jej prípadnou nekonvexnosťou.

Tabuľka 4.1: Rozdelenie vektora Z

(z_1, z_2)	P (z_1, z_2)	F(z_1, z_2)
(1,1)	0,05	0,05
(1,2)	0,05	0,1
(1,3)	0,05	0,15
(2,1)	0,15	0,2
(2,2)	0,1	0,35
(2,3)	0,15	0,55
(3,1)	0,15	0,35
(3,2)	0,2	0,7
(3,3)	0,1	1



Obr. 4.1: Množina prípustných riešení

4.3 p-level eficientné body

Označíme

$$Z_p = \{z \in R^m | F_Z(z) \geq p\}. \tag{4.4}$$

Pokúsime sa množinu prípustných riešení pre úlohu (4.1) ekvivalentne zapísať. Máme

$$\begin{aligned} M &= \{x \in R^n | F_Z(g(x)) \geq p\} \\ &= \{x \in R^n | g(x) \in Z_p\} \\ &= \{x \in R^n | \exists z \in R^m : F(z) \geq p, g(x) = z\}. \end{aligned} \tag{4.5}$$

Označme

$$M^* = \{x \in R^n | \exists z \in R^m : F(z) \geq p, g(x) \geq z\}.$$

Zrejme platí $M \subseteq M^*$ podľa tretieho vyjadrenia pre M z (4.5).
Nech teraz $x^* \in M^*$. Tzn. našli sme $z^* \in R^m : F_Z(z^*) \geq p, g(x^*) \geq z^*$. F_Z je neklesajúca, takže platí $F(g(x^*)) \geq F(z^*) \geq p$. Tzn. $x^* \in M$ podľa prvého vyjadrenia množiny M z (4.5) a $M^* \subseteq M$. Spolu $M = M^*$.
Podľa predchádzajúceho úloha (4.1) je ekvivalentá s

$$\begin{gathered} \text{Min} \quad c(x) \\ \text{za podmienok} \\ g(x) \in Z_p \\ x \in X. \end{gathered} \tag{4.6}$$

Použitím $M = M^*$ úlohu môžeme ekvivalentne prepísať na tvar

$$\begin{gathered} \text{Min} \quad c(x) \\ \text{za podmienok} \\ g(x) - z \geq 0 \\ F_Z(z) \geq p \\ x \in X, \end{gathered} \tag{4.7}$$

Nasledujúca definícia je prebratá z([3], str.116)

Definícia 4.5. *Nech $p \in (0,1)$ a $Z \in R^m$ je náhodný vektor. Bod $v \in R^m$ sa nazýva p-level eficientný bod (PLEP) distribučnej funkcie F_Z, keď platí $F_Z(v) \geq p$ a keď neexistuje $z \leq v, z \neq v$ také, že $F_Z(z) \geq p$.*

Poznámka 4.6. *Ak má úloha optimálne riešenie potom aspoň jeden z PLEP bodov je súčasťou optimálneho riešenia úlohy (4.7).*

Dôkaz
Ak (x^*, z^*) je optimálnym riešením úlohy potom musí platiť $g(x^*) - z^* \geq 0$. Potom aj pre každé $z^{**} \leq z^*$ je splnené $g(x^*) - z^{**} \geq 0$. Účelová funkcia nie je funkciou náhodného vektora Z a teda hodnota z^* ju nijak neovplyvní. Ak pre z^{**} platí aj $F_Z(z^{**}) \geq p$, potom (x^*, z^{**}) je taktiež optimálnym riešením úlohy. Za z^{**} vezmime nejaký minimálny bod pre ktorý platí $z^{**} \leq z^*$ a $F_Z(z^{**}) \geq p$, a to je PLEP bod podľa definície.

QED.

Keď spočítame všetky PLEP body distribučnej funkcie, úlohu si takto môžeme značne zjednodušiť, pretože už nemusíme ďalej hľadať optimálnu hodnotu pre všetky $z \in R^m$ spĺňajúce $F_Z(z) \geq p$, ale stačí hľadať medzi PLEP bodmi. Označíme $v^j, j \in J$ všetky PLEP body úlohy (4.7). Potom úlohu môžeme prepísať na

$$\begin{gathered} \text{Min} \quad c(x) \\ \text{za podmienok} \\ g(x) - \sum_{j \in J} \lambda_j v^j \geq 0 \\ \sum_{j \in J} \lambda_j = 1 \\ \lambda_j \in \{0, 1\} \quad \forall j \in J \\ x \in X, \end{gathered} \tag{4.8}$$

Z tohoto zápisu je jasné, že úlohu môžeme riešiť pre každý PLEP bod zvlášť. Získavame tak $|J|$ úloh konvexného programovania, ktoré už dokážeme riešiť algoritmami pre konvexné úlohy. Získali by sme tak $|J|$ riešení a vybratím riešenia s minimálnou hodnotou účelovej funkcie by sme získali konečné riešenie celej úlohy. Prakticky tento postup môžeme aplikovať na menšie úlohy, ale pre väčšie úlohy počítať každú úlohu zvlášť vychádza ako časovo veľmi neefektívne. Existujú postupy ako neprechádzať neperspektívne PLEP, ďalším by mohlo byť uvoľnenie obmedzenia $\lambda_j \in \{0, 1\}$ na $\lambda_j \in [0, 1]$.([4], str. 15) Uvoľnená úloha by potom vyzerala nasledovne

$$\begin{gathered} \text{Min} \quad c(x) \\ \text{za podmienok} \\ g(x) - \sum_{j \in J} \lambda_j v_j \geq 0 \\ \sum_{j \in J} \lambda_j = 1 \\ \lambda_j \in [0, 1] \quad \forall j \in J \\ x \in X, \end{gathered} \tag{4.9}$$

Keďže všetky obmedzenia v úlohe (4.9)už vymedzujú množiny konvexné, zíkali by sme tak konvexnú množinu prípustných riešení obsahujúcu pôvodnú množinu a úlohu by sme vyriešili niektorým z algoritmov konvexného programovania. Problémom tejto metódy je, že uvoľnená množina nemusí byť konvexným obalom množiny pôvodnej. Získané optimálne riešenie x^*, nemusí splňovať pravdepodobnostné obmedzenie na úrovni p. Skutočnú pravdepodobnosť s akou riešenie spĺňa podmienky vypočítame zo znalosti distribučnej funkcie $p^* = P(g(x^*) \geq Z) = F(g(x^*))$.

Pre prípad celočíselného programovania tento problém pre uvoľnenú úlohu nenastane v prípade, že $g(x)$ môže nadobúdať iba celočíselné hodnoty a za istého predpokladu α-konkavity pre distribučnú funkciu náhodného vektora Z. Nasledujúca definícia je z ([3], str. 105).

Definícia 4.7. *Povieme, že distribučná funkcia F_Z náhodného vektora Z je α-konkávna na množine $A \subseteq R^m$ s $\alpha \in [-\infty, +\infty]$ keď*

$$F_Z \geq m_\alpha(F_Z(x), F_Z(y), \lambda),$$

pre všetky $x, y, z \in A$ a všetky $\lambda \in (0, 1)$ také, že $z \geq \lambda x + (1 - \lambda) y$.

Nasledujúca veta je z ([3] str. 120).

Veta 4.8. *Nech A je množina všetkých možných hodnôt celočíselného náhodného vektora Z. Nech distribučná funkcia F_Z náhodného vektora Z je α-konkávna na $A + Z_+^m$ pre nejaké $\alpha \in [-\infty, +\infty]$, potom pre každé $p \in (0,1)$ máme*

$$Z_p = \left\{ y \in R^m | y \geq z \geq \sum_{j \in J} \lambda_j v^j, \sum_{j \in J} \lambda_j = 1, \lambda_j \geq 0, z \in Z^m \right\},$$

kde $v^j, j \in J$ sú p-eficientné body F_Z.

Máme rovnosti:

$$M = \{x \in R^n | F_Z(g(x)) \geq p\} = \{x \in R^n | g(x) \in Z_p\}.$$

V prípade, že je splnená podmienka celočíselnosti náhodného vektora Z a F_Z je α-konkávna na $A + Z_+^m$,kde A je množina všetkých možných hodnôt Z potom potom podľa vety 4.8

$$M = \{x \in R^n | g(x) \geq z \geq \sum_{j \in J} \lambda_j v^j, \sum_{j \in J} \lambda_j = 1, \lambda_j \geq 0, z \in Z^m\}.$$

V prípade, že $g(x)$ môže nadobúdať iba celočíselné hodnoty potom

$$M = \{x \in R^n | g(x) \geq \sum_{j \in J} \lambda_j v^j, \sum_{j \in J} \lambda_j = 1, \lambda_j \geq 0\}.$$

Za týchto podmienok je teda úloha (4.1) ekvivalentná úlohe (4.9) t.j. úlohe uvoľnenou množinou prípustných riešení a riešenie úlohy (4.9) bude s istotou splňovať požadované pravdepodobnostné obmedzenia.
Viac o p-level eficientných bodoch a úlohe s uvoľnenou množinou prípustných riešení sa dá dočítať v [4].

4.4 Rovnomerné diskrétne rozdelenie pre Kataokov model

Zanecháme teraz úlohu (4.1) a vrátime sa ku Kataokovmu modelu (2.6)(kapitola 2, str.5). V sekcii 3.1.3 sme si ukázali ekvivalentný prepis Kataokovho modelu s predpokladom normálneho rozdelenia vektora Z na úlohu nelineárneho programovania, ktorú pre $p \geq \frac{1}{2}$ vieme riešiť štandardnými algoritmami konvexného programovania.
Uvažujme teraz rovnomerné diskrétne rozdelenie vektoru Z na konečnej množine $\{Z^1, ..., Z^r\}, Z^i \in R^n, i = 1, ..., r$. $\forall x \geq 0$ platí

$$P(Z^T x \geq d) = \frac{1}{r} \sum_{i=1}^{r} \mathbb{1}_{[(Z^i)^T x \geq d]},$$

kde $\mathbb{1}$ je identifikátorová funkcia

$$\mathbb{1}_{[(Z^i)^T x \geq d]} = 1 \text{ ak } (Z^i)^T x \geq d,$$
$$\mathbb{1}_{[(Z^i)^T x \geq d]} = 0 \text{ ak } (Z^i)^T x < d.$$

Z toho máme: $\forall x \geq 0$ platí

$$P(Z^T x \geq d) \geq p \Leftrightarrow \sum_{i=1}^{r} \mathbb{1}_{[(Z^i)^T x \geq d]} \geq pr.$$

To znamená, že $x \geq 0$ bude splňať pravdepodobnostné obmedzenia práve vtedy, keď bude splnených aspoň $p * r$ obmedzení $(Z^i)^T x \geq d$.
Inak povedané $x \geq 0$ bude splňať pravdepodobnostné obmedzenia práve vtedy, keď bude nesplnených najviac $r - \lceil rp \rceil$ zo všetkých obmedzení $(Z^i)^T x \geq d$, kde $\lceil \cdot \rceil$ je horná celá časť čísla.
Z hore uvedeného môžeme úlohu

$$\begin{gathered} \max d \\ \text{za podmienok} \\ \mathrm{P}(\textstyle\sum_{i=1}^{n} Z_i x_i \geq d) \geq p \\ \textstyle\sum_{i=1}^{n} x_i = M, x \geq 0, \end{gathered}$$

ekvivalentne prepísať na úlohu

$$\begin{gathered} \max d \\ \text{za podmienok} \\ (Z^i)^T x + u_i N \geq d, i = 1, ..., r \\ u_i \in \{0, 1\}, i = 1, ...r \\ \sum_{i=1}^{r} u_i \leq r - \lceil rp \rceil \\ \sum_{i=1}^{n} x_i = M, x \geq 0, \end{gathered} \tag{4.10}$$

kde $N \in R$ je konštanta, ktorú volíme dostatočne veľkú, aby sme si boli istý, že $\forall x \in R^n, \sum_{i=1}^{n} x_i = M, x \geq 0$ platí

$$(Z^i)^T x + N \geq d, i = 1, ...r.$$

takže napríklad $N := \max\{M Z^i_j, i = 1, ..., r, j = 1, ..., n\} + 1$.
To nám zaručí to, že ak $u_i = 1$ potom i-tá nerovnosť je automaticky splnená.
Úloha (4.10) je už štandardnou úlohou lineárneho celočíselného programovania.

Kapitola 5

Empirická aplikácia

V tejto kapitole ilustrujeme použitie Kataokovho modelu (2.6)(kapitola 2, str.5) vyriešením úlohy optimalizácie portfólia na základe reálnych historických dát z Burzy cenných papierov Praha. Úlohu riešime najprv pre diskrétne potom pre normálne rozdelenie vektora Z.

5.1 Formulácia úlohy

Pomocou Kataokovho modelu hľadáme optimálne portfolio zložené z akcií vybraných deviatich spoločností obchodovaných na Pražskej burze, pre ktoré máme dostupné data v sledovanom období. Tieto spoločnosti sú:

- CETV
- ČEZ
- Erste Bank
- Komerční banka
- ORCO
- Pegas Nonwovens
- Philip Morris ČR
- Telefónica O2 CR
- Unipetrol

Data boli pôvodne použité v práci [6], kde bolo vysvetlené ako sa k nim došlo. „Počítali sme s týždennými relativnými výnosmi:

$$R_{t+1} = \frac{P_{t+1} - P_t}{P_t},$$

kde R_{t+1} je relativný výnos v čase $t+1$, P_t je cena akcie v čase t. U spoločností ČEZ, Erste Bank, Komerční banka, ORCO, Pegas Nonwovens, Philip Morris ČR, Telefónica O2 CR a Unipetrol sme data ďalej upravili o vyplatené dividendy D_t:

$$R_{t+1} = \frac{P_{t+1} - P_t + D_t}{P_t}.$$

Týždenné data sú od 1.4.2007 do 31.3.2012.“ Našou úlohou je nájsť rozdelenie kapitálu do jednotlivých možností spolu s najväčšou možnou hodnotou $d \in R$ takou, že týždenný relatívny výnos portfólia bude aspoň d s pravdepodobnosťou aspoň p. Problém nadobúda formy (2.6)(kapitola 2, str.5), pre prehľadnosť úlohu pripomenieme.

$$\begin{gathered} \max d \\ \text{za podmienok} \\ \mathrm{P}(\textstyle\sum_{i=1}^{n} Z_i x_i \geq d) \geq p \\ \textstyle\sum_{i=1}^{n} x_i = M, x \geq 0, \end{gathered}$$

kde v tomto prípade rozhodovací vektor bude $(x_1, ..., x_9, d)$, M bude konštanta $M = 1$. Náhodný vektor $Z = Z_1, ..., Z_9$ budú relatívne výnosy v ďalšom týždni. Úlohu vyriešime s predpokladom normálneho a potom diskrétneho rozdelenia pre náhodný vektor Z. Úlohu riešime v programe GAMS, testy a odhady robíme v programe R.

5.2 Výsledky pre diskrétne rozdelenie

Predpokladajme, že historické data dobre popisujú diskrétne rozdelenie náhodného vektora Z. Pre empirické rozdelenie pre Z odvodené z dát teraz vyriešime úlohu Katakovho modelu pre $p = 0,8$, $p = 0,9$, $p = 0,99$. Úlohu vyriešime pomocou ekvivalentného prepisu na úlohu celočíselného lineárneho programovania (4.10)(kapitola 4, str.33) Výsledky sú v tabuľke 5.1.

Tabuľka 5.1: Výsledky pre diskrétne rozdelenie

p	0.8	0.9	0.99
CETV	0	0	0
ČEZ	0.241	0.015	0
Erste Bank	0	0.003	0
Komerční banka	0.068	0.172	0
ORCO	0	0	0
Pegas Nonwovens	0.025	0.057	0.28
Philip Morris ČR	0.228	0.151	0.105
Telefónica O2 CR	0.379	0.526	0.615
Unipetrol	0.058	0.076	0
d	-0.009	-0.026	-0.049

Z tabuľky 5.1 vidíme, že bez ohľadu na zvolené riziko p portfólia nebudú obsahovať vôbec alebo skoro vôbec tituly CETV, Erste Bank a ORCO, ktoré boli v doterajšom období najviac stratové. S rastúcim p presúvame stále väčší podiel do najmenej rizikového aktíva Telefónica O2 CR.
Interpretujme teraz výsledok pre zvolené riziko $p = 0,9$. Vyriešením úlohy pre diskrétne rozdelenie sme našli maximálny možný relatívny výnos $d = -0.026$ taký, že existuje portfólio, z ktorého dosiahneme relatívny výnos aspoň $d = -0.026$ s pravdepodobnosťou aspoň $0,9$. Toto portfólio by vyzeralo takto: 1,5% do akcií ČEZ, 0,3% do akcií Erste Bank, 17,2% do akcií Komerční banky 5,7%

do akcií Pegas Nonwovens, 15,1% do akcií Philip Morris ČR, 52,6% do akcií Telefónica O2 CR a 7,6% do akcií Unipetrol.

5.3 Výsledky pre normálne rozdelenie

Shapirov-Wilkov test pre mnohorozmerné normálne rozdelenie normalitu zamieta s p-hodnotou rádovo 10^{-16}. Normalitu dát v tomto prípade zamietneme na hladine $\alpha = 0,05$. Aj napriek porušenej normalite uvedieme pre porovnanie výsledky pre normálne rozdelenie vektora Z.
Máme teda úlohu s predpokladom normálneho rozdelenia pre Z. O dátach predpokladáme, že sú náhodným výberom pochádzajúcim z 9-rozmerného normálneho rozdelenia $N(\mu, C)$. Ako nestranný odhad pre vektor μ použijeme výberový priemer(tabuľka 5.2). Ako nestranný odhad pre kovariančnú maticu C použijeme výberovú kovariančnú maticu(tabuľka 5.3). Úlohu potom riešime ako úlohu nelineárneho programovania pomocou ekvivalentného prepisu Kataokovho modelu na úlohu konvexného nelineárneho programovania(3.4)(kapitola 3, str.15).

Tabuľka 5.2: Nestranný odhad μ

$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_5$	$\hat{\mu}_6$	$\hat{\mu}_7$	$\hat{\mu}_8$	$\hat{\mu}_9$
-0.0056	0.0010	-0.0018	0.0024	-0.0096	-0.00008	0.0034	0.0010	0.0002

Tabuľka 5.3: Nestranný odhad C

0.0089	0.0023	0.0043	0.0027	0.0047	0.0013	0.0007	0.0009	0.0026
0.0023	0.0015	0.0015	0.0011	0.0015	0.0006	0.0004	0.0005	0.0012
0.0043	0.0015	0.0057	0.0025	0.0030	0.0011	0.0009	0.0007	0.0019
0.0027	0.0011	0.0025	0.0028	0.0015	0.0008	0.0006	0.0006	0.0012
0.0047	0.0015	0.0030	0.0015	0.0079	0.0015	0.0008	0.0009	0.0019
0.0013	0.0006	0.0011	0.0008	0.0015	0.0015	0.0006	0.0005	0.0009
0.0007	0.0004	0.0009	0.0006	0.0008	0.0006	0.0013	0.0003	0.0004
0.0009	0.0005	0.0007	0.0006	0.0009	0.0005	0.0003	0.0007	0.0006
0.0026	0.0012	0.0019	0.0012	0.0019	0.0009	0.0004	0.0006	0.0024

Z tabuľky 5.3 vidíme, že bez ohľadu na zvolené riziko p portfólia budú obsahovať iba tituly ČEZ, Pegas Nonwovens, Philip Morris ČR a Telefónica O2 CR. Porfólia sú si dosť podobné, no s rastúcim p máme vôľu investovať väčší podiel do titulu Pegas Nonwovens a to hlavne na úkor titulu Philip Morris ČR.
Interpretujme teraz výsledok pre zvolené riziko $p = 0,9$. Za predpokladu normálneho rozdelenia náhodného vektora sme našli maximálny možný relatívny výnos $d = -0,02922$ taký, že existuje portfólio, z ktorého dosiahneme relatívny výnos aspoň $d = -0,02922$ s pravdepodobnosťou aspoň $0,9$. Toto portfólio by vyzeralo takto:

Tabuľka 5.4: Výsledky pre normálne rozdelenie

p	0.8	0.9	0.99
CETV	0	0	0
ČEZ	0.104	0.103	0.103
Erste Bank	0	0	0
Komerční banka	0	0	0
ORCO	0	0	0
Pegas Nonwovens	0.014	0.032	0.046
Philip Morris ČR	0.311	0.287	0.267
Telefónica O2 CR	0.572	0.578	0.584
Unipetrol	0	0	0
d	-0.01861	-0.02922	-0.05436

10,3% do akcií ČEZ, 3,2% do akcií Pegas Nonwovens, 28,7% do akcií Philip Morris ČR, 57,8% do akcií Telefónica O2 CR.

5.4 Porovnanie výsledkov

Porovnáme výsledky pre diskrétne a normálne rozdelenie. Ani u normálneho ani u diskrétneho rozdelenia sa v navrhovanom portfóliu pre žiadne p nevyskytli tituly CETV a ORCO. Titul Erste bank sa v portfóliu vyskytol iba raz a to u diskrétneho rozdelenia pre $p = 0,9$ s malým podielom 0,3%.
Pre menšie p, $p = 0,8$,$p = 0,9$ dostávame u normálneho rozdelenia menšiu diverzifikáciu portfólia ako u rozdelenia diskrétneho. U normálneho rozdelenia pre menšie p vôbec neinvestujeme do titulov Komerční banka ani Unipetrol, kým u diskrétneho rozdelenia investujeme.
Pre $p = 0,99$ je to naopak a u normálneho rozdelenia dostávame väčšiu diverzifikáciu portfólia ako u rozdelenia diskrétneho. Vtedy, na rozdiel od navrhovaného portfólia pre normálne rozdelenie, u diskrétneho rozdelenia vôbec neinvestujeme do titulu ČEZ.

Kapitola 6

Záver

V našej práci sme sa zaoberali problematikou stochastického programovania a konvexity úloh s pravdepodobnostnými obmedzeniami. V kapitole 2 sme si formulovali niektoré modely stochastického programovania a ukázali sme si možnosti ich využitia. V druhej kapitole sme sa venovali postačujúcim podmienkam pre konvexitu úloh s pravdepodobnostnými obmedzeniami pre spojité rozdelenia. Zásadný význam pre nás mali veta 3.15 a vety 3.7,3.12, ktoré sme dokázali. Na základe týchto viet sme potom odvodili postačujúce podmienky konvexity pre základnú úlohu (2.3) a pre úlohu (2.5) pre prípad náhodnej pravej strany v stochastických obmedzeniach. Pre prípad náhodnej matice v stochastických obmedzeniach v úlohe (2.5) sme si uviedli výsledky pre združené normálne rozdelenie a združené log-normálne rozdelenie náhodných prvkov v matici. Prípadu log-normálneho rozdelenia sa týkala veta 3.20, ktorú sme dokázali. Úloha Kataokovho modelu (2.6) sa za predpokladu n-rozmerného normálneho rozdelenia ukázala byť konvexná, keď volíme $p \geq \frac{1}{2}$. Uviedli sme aj postačujúcu podmienku pre to, aby obmedzenia podmienených stredných hodnôt vytyčovali konvexnú množinu.
Ďalej sme zaviedli kalkulus pre α-konkávne miery a funkcie, ktorý se potom použili k dôkazom log-konkavity niektorých hustôt a dôkazom konvexity konkrétnych množín zadaných stochastickými obmedzeniami.
U diskrétnych rozdelení sme sa zamerali na úlohu (4.1), kde náhodná je iba pravá strana nerovnosti v stochastických obmedzeniach. Predstavili sme si postačujúce podmienky pre konvexitu úlohy a dokázali vlastnosť log-konkavity pre postupnosti pre niektoré diskrétne rozdelenia. Na jednoduchom príklade sme si ilustrovali problémy, ktoré môžu vzniknúť v súvislosti s konvexitou množiny prípustných riešení u diskrétnych rozdelení. Na riešenie úloh s diskrétnym rozdelením sme si predstavili techniku p-level eficientných bodov, pomocou ktorej dokážeme vyriešiť úlohu s diskrétnym rozdelením aj bez splneného predpokladu konvexity.
V piatej kapitole sme vyriešili úlohu optimalizácie portfólia pomocou Kataokovho modelu na základe reálnych historických dát s predpokladom diskrétneho, potom normálneho rozdelenia výnosov. Pre rovnomerné diskrétne rozdelenie sme úlohu vyriešili pomocou ekvivalentného prepisu úlohy na úlohu lineárneho celočíselného programovania, ktorý sme si ukázali v kapitole pre diskrétne rozdelenia. Pre spojité viacrozmerné normálne rozdelenie sme úlohu vyriešili pomocou prepisu úlohy na nelineárnu úlohu, ktorý sme ukázali v kapitole pre spojité rozdelenia. Výsledky sme potom interpretovali.

Literatúra

[1] P. Lachout, *Matematické programování* - pracovní text k prednášce ,,EKN011 Optimalizace I"zo dňa 26.10.2008

[2] A. Ruszczyński, A. Shapiro *Stochastic programming*- Handbooks in Operation Research and Management Science, Volume 10, Elsevier, Amsterdam 2003

[3] A. Shapiro,D. Dentcheva,A Ruszczyński, *Lectures on stochastic programming*, Modeling and Theory, MPS-SIAM Series on Optimization, Philadelhia 2009

[4] K. Murgaš, *Úlohy pravděpodobnostního programování s diskrétním rozdělením*, Diplomová práce, MFF UK, 2010

[5] A. Prékopa, *On logarithmic concave measures and functions*, Acta Scientiarum Mathematicarum, 34, 335-343, 1973

[6] L. Klouda, *Semi-infinitní programování: teorie a aplikace na eficienci portfolia*, Diplomová práce, MFF UK, 2012