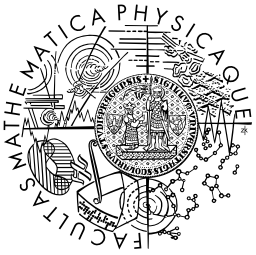


MATEMATICKO-FYZIKÁLNÍ
FAKULTA
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Miloslav Drobný

Optimalizační úlohy s pravděpodobnostními omezeními

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce:	Mgr. Lukáš Adam, Ph.D.
Studijní program:	Matematika
Studijní obor:	Finanční a pojistná matematika

Praha 2018

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) 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 má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 11.5.2018 Podpis autora

V této části bych chtěl poděkovat všem, kteří mi pomáhali a podporovali mě při přípravě a samotném sepisování diplomové práce. V první řadě bych rád poděkoval svému vedoucímu Mgr. Lukáši Adamovi, Ph.D. za jeho trpělivost a obrovskou vstřícnost během celého procesu vytváření práce. Moje další poděkování míří k rodině, zejména pak k manželce, která mě po celou dobu podporovala a vždy se snažila pro mě vytvořit takové podmínky, abych mohl zdárně dokončit sepisování diplomové práce a celé studium. Dále bych chtěl poděkovat všem kolegům na pracovišti MR ZRIR v IKEM, především Ing. Milanu Hájkovi, DrSc. a Mgr. Petru Šedivému, bez jejichž podpory bych toto studium vůbec nemohl absolvovat. Také bych jim chtěl poděkovat za poskytnutí dat, která byla v rámci diplomové práce zpracovávána. V neposlední řadě nesmím opomenout konzultanta RNDr. Martina Brandu, Ph.D, jemuž děkuji za podnětné rady a doporučení.

Název práce: Optimalizační úlohy s pravděpodobnostními omezeními

Autor: Miloslav Drobný

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Mgr. Lukáš Adam, Ph.D., Southern University of Science and Technology, Shenzhen, China. Ústav teorie informace a automatizace AV ČR, Praha, Česká republika

Abstrakt: Autor se v diplomové práci zabývá optimalizačními úlohami s pravděpodobnostními omezeními. Konkrétně pak situacemi, kdy není známo pravděpodobnostní rozdělení přítomného náhodného efektu. K řešení těchto problémů lze přistoupit metodami *optimistických* a *pesimistických scénářů*, kdy z dané rodiny možných pravděpodobnostních rozdělení vybíráme buď nejpříznivější možnou variantu, nebo naopak tu nejméně výhodnou. Optimalizační úlohy s pravděpodobnostními omezeními formulovanými pomocí výše zmíněných přístupů byly za učinění jistých předpokladů transformovány do jednodušších a řešitelných optimalizačních úloh. Dosažené výsledky byly aplikovány na reálná data z oblastí *optimalizace portfolia* a *zpracování obrazu*.

Klíčová slova: pravděpodobnostní omezení, optimalizace portfolia, zpracování obrazu

Title: Optimization problems with chance constraints

Author: Miloslav Drobný

Department: Department of Probability and Mathematical Statistics

Supervisor: Mgr. Lukáš Adam, Ph.D., Southern University of Science and Technology, Shenzhen, China. Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic

Abstract: The diploma thesis is focused on optimization problems with chance constraints, specifically, on the cases when the probability distribution is not known. Two possible approaches are the *pessimistic* and *optimistic scenarios*, where the most favorable and the most unfavorable probability distribution from a given distribution family set is used. Under certain assumptions, we transformed these problems into simpler and easily tractable ones. These stochastic optimalization tasks were transformed into easier and solvable problems. The obtained results were apllied to the real data from *portfolio optimalization* and *image reconstruction* branches.

Keywords: chance constraints, portfolio optimalization, image reconstruction

Obsah

Úvod **2**

1 Základní pojmy a definice **3**

2 Stochastické programování **7**
2.1 Náhodný efekt v účelové funkci 7
2.2 Pravděpodobnostní omezení 8

3 Pravděpodobnostní omezení bez znalosti daného rozdělení **11**
3.1 Pesimistický scénář 11
3.2 Ověření předpokladů 21
3.3 Optimistický scénář 25

4 Aplikační část **26**
4.1 Optimalizace portfolia 26
4.2 Zpracování obrazu 29

5 Výpočetní část **34**
5.1 Optimalizace portfolia 34
5.2 Zpracování obrazu 38

Závěr **41**

Literatura **43**

Úvod

V předložené práci se autor zabývá optimalizačními úlohami s pravděpodobnostními omezeními. Konkrétně se jedná o optimalizační úlohy, ve kterých vystupuje náhodný efekt, jež není popsán konkrétním pravděpodobnostním rozdělením, ale pouze rodinou pravděpodobnostních rozdělení.

Práce je rozdělena do pěti hlavních kapitol, ve kterých jsou nejprve nastíněny teoretické základy, jež slouží pro vybudování a odvození přístupů aplikovaných na popsaný problém. První kapitola je věnována uvedení základních pojmů, definic a lemmat, která jsou používána v dalším textu. Následuje část věnovaná *stochastickému programování*, ve které jsou uvedeny základní přístupy a výsledky používané v případě problémů, kdy je pravděpodobnostní rozdělení náhodného efektu známo.

Stěžejní teoretická část práce je obsahem třetí kapitoly, kde je již veškerá pozornost zaměřena na optimalizační úlohy s pravděpodobnostními omezeními bez znalosti konkrétního pravděpodobnostního rozdělení náhodného efektu. Jsou zde uvedeny dva možné přístupy řešení obdobných úloh. Jedná se o metodu *pesimistických scénářů*, kdy je z množiny možných pravděpodobnostních rozdělení vybráno to nejméně příznivé. Druhý přístup *optimistických scénářů* spočívá v tom, že je z množiny pravděpodobnostních rozdělení naopak vybráno to nejvýhodnější vzhledem ke splnění pravděpodobnostního omezení. V této kapitole je uvedeno několik vět spolu s důkazy, pomocí kterých je možné za jistých předpokladů tyto úlohy přetransformovat na úlohy jednodušší a řešitelné.

V posledních dvou kapitolách jsou výsledky odvozené v teoretické části aplikovány na reálné praktické problémy z oblastí *optimalizace portfolia* a *zpracování obrazu*. V prvním kroku jsou zkoumané problémy transformovány do formy optimalizačních úloh v požadovaném tvaru, na které pak může být použit aparát vybudovaný v předchozích částech. Takto připravená metoda je aplikována na reálná data. V jednom případě se jedná o portfolio tvořené souborem vybraných akcií a v druhém jde o odstraňování šumu z obrázku naměřeného magnetickou rezonancí.

1. Základní pojmy a definice

V této části jsou uvedeny základní definice a lemmata, která jsou spojena s optimalizačními problémy a na která bude odkazováno v dalších kapitolách. Nejprve zaveďme obecnou optimalizační (minimalizační) úlohu, která se nejčastěji zapisuje ve tvaru:

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ s.t. \quad & g_i(\boldsymbol{x}) \leq 0,\ i = 1,...,m \\ & h_j(\boldsymbol{x}) = 0,\ j = 1,...,k \\ & \boldsymbol{x} \in \mathbb{R}^n, \end{aligned} \tag{1.1}$$

kde $\boldsymbol{x}$ je *rozhodovací proměnná* a $f(\boldsymbol{x})$ je *účelová funkce*. Množina, ze které jsou vybírána možná $\boldsymbol{x}$, je restringována m omezeními ve tvaru nerovností a k omezeními ve tvaru rovností, jež dohromady tvoří množinu přípustných řešení. Účelovou funkci $f(\boldsymbol{x})$ (a samozřejmě i všechny ostatní) použitou ve formulaci optimalizační úlohy (1.1) chápeme jako jednorozměrnou (tj. $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$). V opačném případě by se jednalo o problém vícekriteriální optimalizace, kterým se v tomto textu nebudeme zabývat.

Vítanou vlastností optimalizačních úloh je konvexita, které je věnováno několik dalších definic a lemmat.

Definice 1.1 (Konvexita)**.** Řekneme, že

- množina A je konvexní, pokud $\forall a,\ b \in A$ a $\forall \lambda \in [0,1]$ platí: $\lambda a + (1-\lambda)b \in A$.
- funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je konvexní na množině $D \subseteq \mathbb{R}^n$, jestliže $\forall \boldsymbol{x},\ \boldsymbol{y} \in D$ a $\forall \lambda \in [0,1]$ platí: $f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$.
- funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je ryze konvexní na množině $D \subseteq \mathbb{R}^n$, jestliže $\forall \boldsymbol{x},\ \boldsymbol{y} \in D : \boldsymbol{x} \neq \boldsymbol{y}$ a $\forall \lambda \in [0,1]$ platí: $f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) < \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$.
- funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je konkávní na množině $D \subseteq \mathbb{R}^n$, jestliže je funkce $-f(\boldsymbol{x})$ konvexní na stejné množině.

Definice 1.2 (Kvazikonkavita)**.** Řekneme, že funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je kvazikonkávní na množině $D \subseteq \mathbb{R}^n$, jestliže $\forall \boldsymbol{x},\ \boldsymbol{y} \in D$ a $\forall \lambda \in [0,1]$ platí: $f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq \min\{f(\boldsymbol{x}), f(\boldsymbol{y})\}$.

Lemma 1.3. *Nechť je funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ konkávní na množině $D \subseteq \mathbb{R}^n$. Potom je také kvazikonkávní na stejné množině.*

Důkaz. Nechť $\boldsymbol{x}_1,\ \boldsymbol{x}_2 \in D \subseteq \mathbb{R}^n$ a $\lambda \in [0,1]$. Z konkavity funkce $f(\boldsymbol{x})$ na množině D vyplývá:

$$\begin{aligned} & f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \geq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2) \\ & \geq \lambda \min\{f(\boldsymbol{x}_1), f(\boldsymbol{x}_2)\} + (1-\lambda) \min\{f(\boldsymbol{x}_1), f(\boldsymbol{x}_2)\} = \min\{f(\boldsymbol{x}_1), f(\boldsymbol{x}_2)\}. \end{aligned}$$

□

Úloha (1.1) je konvexní, pokud je konvexní jak množina přípustných řešení, tak i účelová funkce $f(\boldsymbol{x})$. Množina přípustných řešení bude pro takto definovanou úlohu konvexní, pokud funkce $g_i(\boldsymbol{x})$ vystupující v omezeních ve tvaru nerovností budou konvexní a funkce $h_j(\boldsymbol{x})$ definující omezení ve tvaru rovností budou afinní. V případě konvexního problému je zaručeno, že každý bod lokálního minima je současně i bodem minima globálního.

Dalšími důležitými pojmy, které by měly být zmíněny ve spojitosti s optimalizačními úlohami, jsou *Lagrangeova funkce* a *Slaterův bod*. Pro úlohu (1.1) má *Lagrangeova funkce* tvar:

$$L(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} u_i g_i(\boldsymbol{x}) + \sum_{j=1}^{k} v_j h_j(\boldsymbol{x}), \tag{1.2}$$

kde koeficienty u_i, $v_j \in \mathbb{R}$ se nazývají *Lagrangeovy multiplikátory*. Existence *Slaterova bodu* pro konvexní úlohu zaručuje, že primární i duální úloha dosahují stejné optimální hodnoty účelové funkce (*silná dualita*) (viz Kapitola 5.2.3 v knize Boyd a Vandenberghe (2004)).

Definice 1.4 (Slaterův bod)**.** Řekneme, že $\boldsymbol{x'} \in \mathbb{R}^n$ je *Slaterovým bodem* konvexní úlohy (1.1), jestliže $\boldsymbol{x'}$ je přípustným řešením a současně $g_i(\boldsymbol{x'}) < 0 \ \forall i \in \{1,...,m\}$ s výjimkou afinních omezeních.

Věta 1.5. *Nechť $\boldsymbol{x}^* \in \mathbb{R}^n$ je lokálním minimem úlohy (1.1), $\boldsymbol{B}(\boldsymbol{x}) = \{i \in \{1,...,m\} : g_i(\boldsymbol{x}) = 0\}$ je množin indexů aktivních omezeních a platí alespoň jedno z následujících:*

- $\nabla_{\boldsymbol{x}} g_i(\boldsymbol{x}^*) \ \forall i \in \boldsymbol{B}(\boldsymbol{x}^*)$ *a* $\nabla_{\boldsymbol{x}} h_j(\boldsymbol{x}^*) \ \forall j \in \{1,...,k\}$ *jsou lineárně nezávislé,*
- *úloha (1.1) je konvexní a existuje Slaterův bod.*

Potom je $\boldsymbol{x}^*$ stacionárním bodem této úlohy a existuje dvojice $(\boldsymbol{u}^*,\boldsymbol{v}^*)$ splňující tzv. First Order Sufficient Condition (FOSC):

- *přípustnost: $g_i(\boldsymbol{x}^*) \leq 0 \ \forall i$, $h_j(\boldsymbol{x}^*) = 0 \ \forall j$*
- *komplementarita: $u_i^* \geq 0 \ \forall i$, $\sum_{i=1}^{m} u_i^* g_i(\boldsymbol{x}^*) = 0$*
- *optimalita: $\nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*,\boldsymbol{u}^*,\boldsymbol{v}^*) = 0$.*

Další část bude věnována objektům, které jsou spojeny s řešením optimalizačních úloh, nebo pomocí nichž mohou být přímo zadávány množiny přípustných řešení. Jedná se o *kužele* a *epigrafy* funkcí.

Definice 1.6 (Kužele)**.**

- Množina $A \subset \mathbb{R}^n$ se nazývá kužel (s vrcholem v počátku), jestliže $\boldsymbol{0} \in A$ a pro každý bod $\boldsymbol{s} \in A$ a $\alpha > 0$ je $\alpha \boldsymbol{s} \in A$.
- Množina $A \subset \mathbb{R}^n$ se nazývá konvexní kužel, pokud je kuželem a zároveň konvexní množinou.
- Nechť $A \subset \mathbb{R}^n$ je kužel. Potom duálním kuželem ke kuželi A rozumíme množinu $A^* = \{\boldsymbol{y} \in \mathbb{R}^n : \langle \boldsymbol{x},\boldsymbol{y} \rangle \geq 0 \ \forall \boldsymbol{x} \in A\}$.

- Nechť $A \subset \mathbb{R}^n$ je kužel. Potom polárním kuželem ke kuželi A rozumíme množinu $P = -A^* = \{\boldsymbol{y} \in \mathbb{R}^n : \langle \boldsymbol{x},\boldsymbol{y} \rangle \leq 0 \ \forall \boldsymbol{x} \in A\}$.

Lemma 1.7. *Nechť $A \subset \mathbb{R}^n$ je kužel, potom A^* je konvexní uzavřený kužel.*

Důkaz. Nejprve ukážeme, že A^* je kužel. Vezměme tedy $\boldsymbol{l} \in A^*$ a $\alpha > 0$. Potom musí nutně $\forall \boldsymbol{s} \in A$ platit, že $\langle \alpha \boldsymbol{l},\boldsymbol{s} \rangle = \alpha \langle \boldsymbol{l},\boldsymbol{s} \rangle \geq 0$. Tím je dokázáno, že množina A^* je kužel.

Dalším krokem je ukázat, že množina A^* je konvexní. Nechť $\boldsymbol{l}_1$, $\boldsymbol{l}_2 \in A^*$, $\lambda \in [0,1]$, potom $\forall \boldsymbol{s} \in A$ platí, že $\langle \lambda \boldsymbol{l}_1 + (1-\lambda)\boldsymbol{l}_2,\boldsymbol{s} \rangle = \lambda \langle \boldsymbol{l}_1,\boldsymbol{s} \rangle + (1-\lambda)\langle \boldsymbol{l}_2,\boldsymbol{s} \rangle \geq 0$, neboť se jedná o součet dvou nezáporných členů. Tím je dokázáno, že konvexní kombinace prvků z množiny A^* zůstává ve stejné množině ($\lambda \boldsymbol{l}_1 + (1-\lambda)\boldsymbol{l}_2 \in A^*$). To znamená, že množina A^* je konvexní.

Nyní zbývá dokázat, že množina A^* je uzavřená. Pro tento účel vytvořme posloupnost $\boldsymbol{l}_k \in A^* \ \forall k \in \mathbb{N}$, jež konverguje k prvku $\boldsymbol{l}$ ($\boldsymbol{l}_k \to \boldsymbol{l}$ pro $k \to +\infty$). Předpokládejme, že $\boldsymbol{l} \notin A^*$, tudíž $\exists \delta > 0$ a $\boldsymbol{s} \in A : \langle \boldsymbol{l},\boldsymbol{s} \rangle < -\delta$. Ze spojitosti skalárního součinu vyplývá, že musí $\exists k' \in \mathbb{N} : \langle \boldsymbol{l}_k,s \rangle < 0 \ \forall k \geq k'$. To je ovšem v rozporu s výchozím předpokladem, že $\boldsymbol{l}_k \in A^* \ \forall k \in \mathbb{N}$. Tím je dokázáno, že limita každé posloupnosti definované na množině A^* je prvkem množiny A^*, a tudíž je daná množina uzavřená. □

Definice 1.8 (Epigraf)**.** Epigrafem funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ nazveme množinu $\operatorname{epi} f = \{(\boldsymbol{x},\eta) \in \mathbb{R}^n \times \mathbb{R} : f(\boldsymbol{x}) \leq \eta\}$.

V souvislosti s *epigrafem* funkce $f(\boldsymbol{x})$ se zavádí tzv. *úrovňové množiny*, kdy pro každé pevné $\eta \in \mathbb{R}$ má daná množina tvar $\{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) \leq \eta\}$. Pro ujasnění rozdílu mezi těmito dvěma pojmy zdůrazněme, že pro funkci $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je množina $\operatorname{epi} f \subset \mathbb{R}^{n+1}$, zatímco příslušné úrovňové množiny patří do prostoru $\mathbb{R}^n$. Dá se také jednoduše ukázat, že *úrovňové množiny* kvazikonvexních funkcí jsou konvexní množiny.

Na závěr první kapitoly ještě uveďme několik nezbytných definic a lemmat, jejichž znalost a výstupy budou dále využívány.

Definice 1.9 (Charakteristická a indikátorová funkce množiny)**.** Charakteristická funkce množiny A je definována jako:

$$\mathbb{I}_{[x \in A]} = \begin{cases} 0 & \text{pro } x \notin A \\ 1 & \text{pro } x \in A. \end{cases}$$

Indikátorová funkce množiny A je dána předpisem:

$$\delta_A(x) = \begin{cases} 0 & \text{pro } x \in A \\ +\infty & \text{pro } x \notin A. \end{cases}$$

Definice 1.10 (Opěrná funkce)**.** Opěrná funkce množiny A v bodě x je určena vztahem:

$$\sigma_A(x) = \sup_{a \in A} \langle x,a \rangle.$$

Lemma 1.11. *Nechť $P \subset \mathbb{R}^n$ je polární kužel k uzavřenému konvexnímu kuželi A, potom platí:*

$$\sigma_A(\boldsymbol{x}) = \delta_P(\boldsymbol{x}).$$

Důkaz. Nejprve ukážeme, že hodnota opěrné funkce pro libovolný kužel $A \subset \mathbb{R}^n$ nabývá pouze hodnot 0, nebo $+\infty$. Každý kužel obsahuje nulový prvek, a tedy hodnota opěrné funkce nemůže být v žádném bodě záporná ($\langle \boldsymbol{x},\boldsymbol{0}\rangle = 0\ \forall \boldsymbol{x} \in \mathbb{R}^n$). Pokud však existuje prvek $\boldsymbol{a} \in A$ takový, že $\langle \boldsymbol{x},\boldsymbol{a}\rangle > 0$, potom z vlastnosti kužele plyne, že pro libovolné $\alpha > 0$ platí $\alpha\boldsymbol{a} \in A$. Pro $\alpha \to +\infty$ pak dostáváme, že $\sigma_A(\boldsymbol{x}) \geq \alpha\langle \boldsymbol{x},\boldsymbol{a}\rangle \to +\infty$. A tedy supremum je z definice rovno $+\infty$. Proto opěrná funkce kužele může nabývat pouze hodnot 0 nebo $+\infty$.

Nechť $\sigma_A(\boldsymbol{x}) = \sup_{\boldsymbol{a}\in A} \langle \boldsymbol{x},\boldsymbol{a}\rangle = 0$, tudíž $\forall \boldsymbol{a} \in A$ platí $\langle \boldsymbol{x},\boldsymbol{a}\rangle \leq 0$. To znamená, že $\boldsymbol{x} \in P$ ($\delta_P(\boldsymbol{x}) = 0$). Jestliže naopak $\sigma_A(\boldsymbol{x}) = \sup_{\boldsymbol{a}\in A} \langle \boldsymbol{x},\boldsymbol{a}\rangle = +\infty$, tak existuje $\boldsymbol{a} \in A$, pro které platí $\langle \boldsymbol{x},\boldsymbol{a}\rangle > 0$, z čehož vyplývá, že $\boldsymbol{x} \notin P$ ($\delta_P(\boldsymbol{x}) = +\infty$). Tím je daný vztah dokázán. □

Definice 1.12 (Vlastní funkce)**.** Řekneme, že funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je vlastní, jestliže $\exists \boldsymbol{x} \in \mathbb{R}^n$, pro které platí $f(\boldsymbol{x}) < +\infty$, a $f(\boldsymbol{x}) > -\infty\ \forall \boldsymbol{x} \in \mathbb{R}^n$.

Definice 1.13 (Konjugovaná funkce)**.** Konjugovaná funkce k vlastní funkci $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je určena předpisem:

$$f^*(\boldsymbol{\nu}) = \sup_{\boldsymbol{x}\in\mathbb{R}^n} \{\boldsymbol{\nu}^\top \boldsymbol{x} - f(\boldsymbol{x})\}.$$

Pro indikátorovou funkci neprázdné množiny $A \subset \mathbb{R}^n$, jež je z definice funkcí vlastní, platí, že:

$$\delta_A^*(\boldsymbol{x}) = \sup_{\boldsymbol{y}\in\mathbb{R}^n} \{\boldsymbol{x}^\top \boldsymbol{y} - \delta_A(\boldsymbol{y})\} = \sup_{\boldsymbol{y}\in A} \langle \boldsymbol{x},\boldsymbol{y}\rangle = \sigma_A(\boldsymbol{x}). \tag{1.3}$$

Druhá rovnost platí, neboť při výpočtu suprema lze množinu $\mathbb{R}^n$ nahradit množinou A, protože jinak by $-\delta_A(\boldsymbol{y}) = -\infty$, což je varianta, kterou lze vzhledem k uvažovanému supremu opomenout.

Definice 1.14 (Polospojitost)**.** Řekneme, že funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je zdola polospojitá v bodě $\boldsymbol{y} \in \mathbb{R}^n$, jestliže platí:

$$\liminf_{\boldsymbol{x}\to\boldsymbol{y}} f(\boldsymbol{x}) \geq f(\boldsymbol{y}).$$

2. Stochastické programování

V této části bude deterministická úloha (1.1) rozšířena o náhodný efekt $\boldsymbol{\omega} \in \mathbb{R}^l$ s distribuční funkcí $\boldsymbol{F}$. Pomocí tohoto rozšíření je možné se více přiblížit většině reálných situací, ve kterých hraje nejistota obvykle nemalou roli.

2.1 Náhodný efekt v účelové funkci

Nejprve se budeme zabývat situací, kdy se náhodný vektor $\boldsymbol{\omega}$ vyskytuje pouze v účelové funkci. Množina přípustných řešení bude označena jako $\boldsymbol{\mathcal{X}} \subseteq \mathbb{R}^n$ a nebude nijak závislá na náhodném vektoru $\boldsymbol{\omega}$:

$$\min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} f(\boldsymbol{x},\boldsymbol{\omega}). \tag{2.1}$$

Je zřejmé, že úloha definovaná předpisem (2.1) nedává dobrý smysl, neboť pro každou realizaci náhodného vektoru $\boldsymbol{\omega}^*$ bude výběr optimálního řešení $\boldsymbol{x}^*$ odlišný. Jistou cestu řešení nabízí *kritérium střední hodnoty*, kdy je místo úlohy (2.1) řešena úloha:

$$\min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} \int_{\mathbb{R}^l} f(\boldsymbol{x},\boldsymbol{\omega}) \boldsymbol{dF}(\boldsymbol{\omega}). \tag{2.2}$$

Jedná se o minimalizaci střední hodnoty účelové funkce $f(\boldsymbol{x},\boldsymbol{\omega})$ na množině přípustných řešení $\boldsymbol{\mathcal{X}}$. Tento přístup vyžaduje znalost pravděpodobnostního rozdělení náhodného vektoru $\boldsymbol{\omega}$. Další možný postup vychází z *cílového programování*, kdy je optimální řešení hledáno pomocí úlohy:

$$\min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} \int_{\mathbb{R}^l} \mid f(\boldsymbol{x},\boldsymbol{\omega}) - f^*(\boldsymbol{\omega}) \mid \boldsymbol{dF}(\boldsymbol{\omega}), \tag{2.3}$$

kde se minimalizují rozdíly hodnoty účelové funkce od optimální hodnoty při konkrétní realizaci náhodného vektoru $\boldsymbol{\omega}$ vážené jeho pravděpodobnostním rozdělením. Existují samozřejmě i další přístupy. Lze například uvést ještě tyto varianty:

- $\min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} f(\boldsymbol{x},\hat{\boldsymbol{\omega}})$, *použití bodové odhadu* $\hat{\boldsymbol{\omega}}$ *pro náhodnou veličinu* $\boldsymbol{\omega}$
- $\min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} u(f(\boldsymbol{x},\boldsymbol{\omega}))$, *použití vhodné užitkové funkce* u
- $\min_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} \mathbb{E} f(\boldsymbol{x},\boldsymbol{\omega}) + \alpha \mathbb{R}(f(\boldsymbol{x},\boldsymbol{\omega}))$, *zahrnutí rizikové míry* $\mathbb{R}$ *s váhou* α.

Znovu je potřeba zdůraznit, že k použití výše zmíněných přístupů je nutné znát pravděpodobnostní rozdělení náhodného vektoru $\boldsymbol{\omega}$.

2.2 Pravděpodobnostní omezení

Druhou možností, kde se může náhodný efekt v optimalizační úloze objevit, jsou omezující podmínky tvořící množinu přípustných řešení. Taková optimalizační úloha může mít tvar:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ s.t. \quad & g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0,\ i = 1,...,m \\ & \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \subseteq \mathbb{R}^n. \end{aligned} \tag{2.4}$$

V tomto případě nejsou uvažovány omezení ve tvaru rovností, neboť pro náhodné vektory se spojitým pravděpodobnostním rozdělením by tato omezení byla až příliš silná a v některých případech až nesplnitelná. Vzhledem k tomu, že se v omezeních objevuje náhodný vektor $\boldsymbol{\omega}$, není zcela jasné, jak daná omezení chápat. Platnost omezení by mohla být vyžadována *skoro jistě*:

$$\mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0,\ \forall i \in \{1,...,m\}] = 1. \tag{2.5}$$

Tato vždy přípustná řešení však nejsou zcela praktická, protože mohou být až příliš omezující. Mnohdy je zcela postačující, pokud daná omezení platí alespoň s pravděpodobností $\alpha < 1$. Může být tedy požadováno, aby buď všechna omezení najednou platila s pravděpodobností alespoň α, nebo aby každé individuální omezení platilo s předem danou minimální pravděpodobností α_i. První přístup je nazýván jako *sdružené pravděpodobnostní omezení* a druhý jako *individuální pravděpodobnostní omezení*. Pro přesnější formulaci omezujících podmínek optimalizační úlohy (2.4) definujme množinu přípustných řešení pro oba výše zmíněné přístupy:

- *sdružené pravděpodobnostní omezení* $\alpha \in [0,1]$:
  $\boldsymbol{\mathcal{K}}^s(\alpha) = \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0,\ \forall i \in \{1,...,m\}] \geq \alpha\}$
- *individuální pravděpodobnostní omezení* $\boldsymbol{\alpha} = (\alpha_1,...,\alpha_m)$:
  $\boldsymbol{\mathcal{K}}(\boldsymbol{\alpha}) = \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0] \geq \alpha_i,\ \forall i \in \{1,...,m\}\}$
  $= \bigcap_{i=1}^m \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0] \geq \alpha_i\} = \bigcap_{i=1}^m \boldsymbol{\mathcal{K}}_i(\alpha_i).$

Lemma 2.1. *Nechť pravděpodobnostní omezení (2.4) jde vyjádřit ve tvaru $g_i(\boldsymbol{x},\omega_i)$, kde ω_i jsou složky náhodného vektoru $\boldsymbol{\omega}$. Pokud jsou ω_i nezávislé, potom platí:*

$$\boldsymbol{\mathcal{K}}^s(\alpha) = \bigcup_{\boldsymbol{\beta} \in B(\alpha)} \bigcap_{i=1}^m \boldsymbol{\mathcal{K}}_i(\beta_i),$$

kde $B(\alpha)$ je definováno jako množina všech $\boldsymbol{\beta} \in [0,1]^m$ splňujících $\prod_{i=1}^m \beta_i = \alpha$.

Důkaz. Nechť $\boldsymbol{x} \in \boldsymbol{\mathcal{K}}^s(\alpha)$. To znamená, že $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ a všechna pravděpodobnostní omezení jsou současně splněna s pravděpodobností alespoň α. S využitím nezávislosti ω_i lze psát:

$$\mathbb{P}[g_i(\boldsymbol{x},\omega_i) \geq 0,\ \forall i \in \{1,...,m\}] = \prod_{i=1}^m \mathbb{P}[g_i(\boldsymbol{x},\omega_i) \geq 0] \geq \alpha.$$

Nyní položme $\beta_i^* = \mathbb{P}[g_i(\boldsymbol{x},\omega_i) \geq 0]\ \forall i \in \{1,...,m\}$, pro které platí $\prod_{i=1}^m \beta_i^* \geq \alpha$. Následně definujme $\beta_i = \beta_i^*\ \forall i \neq l$, kde $l \in \{1,...,m\}$. β_l volíme tak, aby platilo $\prod_{i=1}^m \beta_i = \alpha$, což znamená, že $\beta_l \leq \beta_l^*$. Z toho plyne, že $\boldsymbol{\beta} \in B(\alpha)$ a také $\boldsymbol{x} \in \bigcup_{\boldsymbol{\beta}\in B(\alpha)} \bigcap_{i=1}^m \mathcal{K}_i(\beta_i)$. Tím je dokázána první inkluze.

V druhém kroku volme $\boldsymbol{x} \in \bigcup_{\boldsymbol{\beta}\in B(\alpha)} \bigcap_{i=1}^m \mathcal{K}_i(\beta_i)$. To znamená, že $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ a $\exists \boldsymbol{\beta} \in [0,1]^m : \prod_{i=1}^m \beta_i = \alpha$. Současně platí, že $\mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0] \geq \beta_i\ \forall i \in \{1,...,m\}$. K tomuto přidejme předpoklad o nezávislosti složek ω_i a celkem dostaneme:

$$\mathbb{P}[g_i(\boldsymbol{x},\omega_i) \geq 0,\ \forall i \in \{1,...,m\}] = \prod_{i=1}^m \mathbb{P}[g_i(\boldsymbol{x},\omega_i) \geq 0] \geq \prod_{i=1}^m \beta_i = \alpha.$$

Z toho plyne, že $\boldsymbol{x} \in \boldsymbol{\mathcal{K}}^s(\alpha)$. Druhá inkluze je tím dokázána, a tedy i rovnost obou množin. □

Věta 2.2. *Nechť funkce $g_i(\boldsymbol{x},\boldsymbol{\omega})$ je separovatelná ve smyslu $g_i(\boldsymbol{x},\boldsymbol{\omega}) = q_i(\boldsymbol{x}) - \omega_i$, $q_i(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ jsou kvazikonkávní a F_i jsou marginální distribuční funkce náhodného vektoru $\boldsymbol{\omega} \in \mathbb{R}^m$, pak je množina $\mathcal{K}_i^*(\alpha_i) = \{\boldsymbol{x} : \mathbb{P}[q_i(\boldsymbol{x}) \geq \omega_i] \geq \alpha_i\}$ konvexní pro libovolné rozdělení náhodného vektoru $\boldsymbol{\omega}$ a libovolné $\alpha_i \in [0,1]$.*

Důkaz. Nerovnost $\mathbb{P}[q_i(\boldsymbol{x}) \geq \omega_i] \geq \alpha_i$ ve formulaci množiny $\mathcal{K}_i^*(\alpha_i)$ lze za použití distribuční funkce F_i náhodné veličiny ω_i přepsat do tvaru $F_i[q_i(\boldsymbol{x})] \geq \alpha_i$. Na obě strany této nerovnosti aplikujeme monotónní transformaci ve formě kvantilové funkce F_i^{-1} náhodné veličiny ω_i. Po přenásobení vzniklé nerovnosti -1 získáme $-q_i(\boldsymbol{x}) \leq -F_i^{-1}(\alpha_i)$. Množina bodů splňující tuto nerovnost tvoří úrovňovou množinu funkce $-q_i(\boldsymbol{x})$. Jelikož je funkce $q_i(\boldsymbol{x})$ dle předpokladů kvazikonkávní, tak je funkce $-q_i(\boldsymbol{x})$ kvazikonvexní, a tudíž jsou její úrovňové množiny konvexní. □

Poznámka 2.3. Z této věty vyplývá, že pro konvexní množinu $\boldsymbol{\mathcal{X}}$ je množina přípustných řešení $\mathcal{K}(\boldsymbol{\alpha}) = \cap_{i=1}^m \mathcal{K}_i^*(\alpha_i) \cap \boldsymbol{\mathcal{X}}$ konvexní, neboť průnik konvexních množin je konvexní množina.

Dále se zaměříme na sdružená pravděpodobnostní omezení. V této souvislosti je potřeba zavést pojem *logaritmické konkavity*.

Definice 2.4 (Logaritmická konkavita)**.** Nechť $\mathbb{P}$ je pravděpodobnostní míra na prostoru $(\mathbb{R}^l,\mathcal{B}^l)$. Řekneme, že:

- $\mathbb{P}$ je log-konkávní, pokud pro libovolné dvě množiny $A,B \in \mathbb{R}^l$ a $\lambda \in [0,1]$ platí: $\mathbb{P}(\lambda A + (1-\lambda)B) \geq \mathbb{P}(A)^\lambda \mathbb{P}(B)^{1-\lambda}$
- funkce $f(\boldsymbol{x}) : \mathbb{R}^l \to \mathbb{R}$ je log-konkávní, jestliže $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^l$ a $\forall \lambda \in [0,1]$ platí: $f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq f(\boldsymbol{x})^\lambda f(\boldsymbol{y})^{1-\lambda}$.

Lemma 2.5. *Nechť náhodný vektor $\boldsymbol{\omega} \sim N_l(\boldsymbol{\mu},\boldsymbol{\Sigma})$, potom hustota pravděpodobnosti $f_{\boldsymbol{\omega}}(\boldsymbol{v})$ je log-konkávní funkce.*

Důkaz. Hustota pravděpodobnosti náhodného vektoru $\boldsymbol{\omega}$ má tvar:

$$f_{\boldsymbol{\omega}}(\boldsymbol{v}) = \frac{1}{\sqrt{(2\pi)^l\,|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{v}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{v}-\boldsymbol{\mu})}.$$

Libovolná funkce je log-konkávní, pokud je její logaritmus konkávní funkcí. Pokud je tedy funkce v exponentu ($f_e(\boldsymbol{v}) = -\frac{1}{2}(\boldsymbol{v}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{v}-\boldsymbol{\mu})$) konkávní, pak

je hustota pravděpodobnosti vícerozměrného normálního rozdělení log-konkávní funkcí. Funkce $f_e(\boldsymbol{v})$ je kvadratická forma tvořená maticí $\boldsymbol{\Sigma}$, která je vzhledem k požadavkům kladených na kovarianční matici pozitivně semidefinitní. Z toho vyplývá, že funkce $f_e(\boldsymbol{v})$ je konkávní a $f_{\boldsymbol{\omega}}(\boldsymbol{v}) \sim \exp\{f_e(\boldsymbol{v})\}$ je log-konkávní. □

Poznámka 2.6. Mezi rozdělení s log-konkávní hustotou pravděpodobnosti patří např. také *gamma rozdělení*, *Dirichletovo rozdělení*, *Wishartovo rozdělení* a další.

Věta 2.7 (Prékopa (1971))**.** *Nechť funkce $g_i(\boldsymbol{x},\boldsymbol{\omega})$, $i = 1,...,m$ jsou konkávní na $\mathbb{R}^n \times \mathbb{R}^l$. Nechť $\boldsymbol{\omega}$ má log-konkávní rozdělení na $(\mathbb{R}^l,\mathcal{B}^l)$, pak je funkce $h(\boldsymbol{x}) = \mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0, \forall i \in \{1,...,m\}]$ log-konkávní na $\mathbb{R}^n$.*

Poznámka 2.8. Pokud je navíc $\boldsymbol{\mathcal{X}}$ konvexní, pak je množina $\boldsymbol{\mathcal{K}}(\alpha) = \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \mathbb{P}[g_i(\boldsymbol{x},\boldsymbol{\omega}) \geq 0, \forall i \in \{1,...,m\}] \geq \alpha\}$ konvexní.

Lemma 2.9. *Nechť funkce $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je log-konkávní, potom je $f(\boldsymbol{x})$ také kvazikonkávní.*

Důkaz. Funkce $f(\boldsymbol{x})$ je log-konkávní, a tudíž $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ a $\lambda \in [0,1]$ platí:

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq f(\boldsymbol{x})^{\lambda} f(\boldsymbol{y})^{1-\lambda} = \left(\frac{f(\boldsymbol{x})}{\min\{f(\boldsymbol{x}),f(\boldsymbol{y})\}}\right)^{\lambda}$$
$$\left(\frac{f(\boldsymbol{y})}{\min\{f(\boldsymbol{x}),f(\boldsymbol{y})\}}\right)^{1-\lambda} \min\{f(\boldsymbol{x}),f(\boldsymbol{y})\} \geq \min\{f(\boldsymbol{x}),f(\boldsymbol{y})\}.$$

Poslední nerovnost platí díky tomu, že oba výrazy v závorkách jsou větší nebo rovny jedné a exponenty jsou nezáporné. □

Věta 2.10. *Nechť jsou funkce $q_i(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ konkávní $\forall i \in \{1,...,m\}$, $\boldsymbol{\omega}$ má log-konkávní rozdělení a množina $\boldsymbol{\mathcal{X}}$ je konvexní. Potom je množina $\boldsymbol{\mathcal{K}}(\alpha) = \boldsymbol{\mathcal{X}} \cap \{\boldsymbol{x} \in \mathbb{R}^n : \mathbb{P}[q_i(\boldsymbol{x}) \geq \omega_i, \forall i \in \{1,...,m\}] \geq \alpha\}$ konvexní.*

Důkaz. Funkce $q_i(\boldsymbol{x})$ jsou konkávní. Potom jsou i funkce $g_i(\boldsymbol{x},\omega_i) = q_i(\boldsymbol{x}) - \omega_i$ konkávní, neboť se jedná o složení konkávní a afinní funkce. Dle Věty 2.7 je funkce $h(\boldsymbol{x}) = \mathbb{P}[g_i(\boldsymbol{x},\omega_i) \geq 0, \forall i \in \{1,...,m\}]$ log-konkávní, a tudíž (viz Lemma 2.9) také kvazikonkávní. Dále platí, že úrovňové množiny kvazikonvexní funkce ($-h(\boldsymbol{x}) \leq -\alpha$) jsou konvexní. Konvexita se zachová i po průniku s konvexní množinou $\boldsymbol{\mathcal{X}}$. □

3. Pravděpodobnostní omezení bez znalosti daného rozdělení

Všechny výsledky odvozené v předchozí kapitole vycházely z předpokladu znalosti pravděpodobnostního rozdělení náhodné veličiny (vektoru) vyskytující se v pravděpodobnostních omezeních. Obvykle je však k dispozici jen minimum informací o daném pravděpodobnostním rozdělení. V některých případech lze alespoň zařadit ono pravděpodobnostní rozdělení $\mathbb{P}$ do určité rodiny pravděpodobnostních rozdělení $\mathcal{P}$, jež může být určena například typem pravděpodobnostních rozdělení či specifikací některých momentů, nebo jiných parametrů definujících pravděpodobnostní rozdělení. Tato situace bude náplní následujícího textu.

3.1 Pesimistický scénář

Pravděpodobnostní omezení za znalosti rodiny pravděpodobnostních rozdělení $\mathcal{P}$ může mít tvar:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{g}(\boldsymbol{x},\boldsymbol{\omega}) \geq 0] \geq 1-\epsilon, \tag{3.1}$$

kde $\epsilon \in [0,1)$ a funkce $\boldsymbol{g} : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^m$ je chápána jako vícerozměrná. Díky tomuto zápisu lze jiným způsobem vyjádřit *sdružené pravděpodobnostní omezení*. Tento tvar reprezentuje tzv. *pesimistický scénář*, neboť z množiny pravděpodobnostních rozdělení $\mathcal{P}$ je vybíráno právě to, pro které je hodnot pravděpodobnosti nejméně příznivá (nejnižší).

Před formulací optimalizační úlohy s omezením ve tvaru *pesimistického scénáře* nejprve zaveďme potřebné definice.

Definice 3.1. Mějme $\mathcal{D} \subseteq \mathbb{R}^n$ konvexní uzavřený kužel, $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbb{R}^n$ a funkci $\boldsymbol{d} : \mathbb{R}^l \to \mathbb{R}^n$. Potom:

- nerovností $\boldsymbol{v} \preceq_{\mathcal{D}} \boldsymbol{w}$ ($\boldsymbol{v} \prec_{\mathcal{D}} \boldsymbol{w}$) rozumíme, že $\boldsymbol{w}-\boldsymbol{v} \in \mathcal{D}$ ($\boldsymbol{w}-\boldsymbol{v} \in \operatorname{int}\mathcal{D}$).
- funkci $\boldsymbol{d}(\boldsymbol{\omega})$ nazveme $\mathcal{D}$-konvexní, pokud $\forall \boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2 \in \mathbb{R}^l$ platí, že $\boldsymbol{d}(\lambda\boldsymbol{\omega}_1 + (1-\lambda)\boldsymbol{\omega}_2) \preceq_{\mathcal{D}} \lambda\boldsymbol{d}(\boldsymbol{\omega}_1) + (1-\lambda)\boldsymbol{d}(\boldsymbol{\omega}_2)$.
- funkci $\boldsymbol{d}(\boldsymbol{\omega})$ nazveme pozitivně homogenní stupně k, pokud $\forall \boldsymbol{\omega} \in \mathbb{R}^l$ a $\alpha > 0$ platí $\boldsymbol{d}(\alpha\boldsymbol{\omega}) = \alpha^k \boldsymbol{d}(\boldsymbol{\omega})$.
- $\mathcal{D}$-epigraf funkce $\boldsymbol{d}(\boldsymbol{\omega})$ je definován předpisem $\operatorname{epi}(\boldsymbol{d}) = \{(\boldsymbol{\omega},\boldsymbol{\eta}) \in \mathbb{R}^l \times \mathbb{R}^d : \boldsymbol{d}(\boldsymbol{\omega}) \preceq_{\mathcal{D}} \boldsymbol{\eta}\}$.

Poznámka 3.2. Pro $\mathcal{D} = \mathbb{R}^n_+$ nerovnost $\boldsymbol{v} \preceq_{\mathcal{D}} \boldsymbol{w}$ znamená, že $v_i \leq w_i \; \forall i \in \{1,\ldots,n\}$.

V této sekci budeme vycházet z optimalizační úlohy s *pesimistickým* pravděpodobnostním omezením (3.1) a speciálním tvarem množiny přípustných řešení:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \subseteq \mathbb{R}^n \\ & \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})] \geq 1-\epsilon, \end{aligned} \tag{3.2}$$

kde $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ je obecná účelová funkce, $\boldsymbol{\mathcal{X}} \subseteq \mathbb{R}^n$ je konvexní polyedrická množina (průnik konečného počtu uzavřených poloprostorů), $\boldsymbol{T}(\boldsymbol{x}) \in \mathbb{R}^{k\times l}$ a $\boldsymbol{u}(\boldsymbol{x}) \in \mathbb{R}^k$ jsou složeny z afinních funkcí proměnné $\boldsymbol{x}$ a $\epsilon \in [0,1)$. Rodina pravděpodobnostních rozdělení $\mathcal{P}$ je definována předpisem:

$$\mathcal{P} = \{\mathbb{P} \in \mathcal{P}_0(\boldsymbol{\Omega}) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\omega}] = \boldsymbol{\mu},\, \mathbb{E}_{\mathbb{P}}[\boldsymbol{d}(\boldsymbol{\omega})] \preceq_{\mathcal{D}} \boldsymbol{\sigma}\}, \tag{3.3}$$

kde $\mathcal{P}_0(\boldsymbol{\Omega})$ je množina Borelovských pravděpodobnostních rozdělení na $\boldsymbol{\Omega} \subseteq \mathbb{R}^l$, $\boldsymbol{\mu} \in \mathbb{R}^l$ je střední hodnota náhodného vektoru $\boldsymbol{\omega}$ a $\boldsymbol{\sigma} \in \mathbb{R}^d$ reprezentuje horní mez pro očekávanou hodnotu disperzní funkce $\boldsymbol{d}(\boldsymbol{\omega}) \in \mathbb{R}^d$. Navíc předpokládejme, že $\mathcal{D} \subseteq \mathbb{R}^d$ je konvexní uzavřený kužel. Dále definujme množinu indexů omezení $\mathcal{I} = \{1,...,k\}$. Pro každé pevné $\boldsymbol{x}$ označme množinu indexů, pro které může být omezení porušeno:

$$\mathcal{I}(\boldsymbol{x}) = \{i \in \mathcal{I} : \boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\omega} \geq u_i(\boldsymbol{x}) \text{ pro nějaké } \boldsymbol{\omega} \in \boldsymbol{\Omega}\} \tag{3.4}$$

Na doplnění ještě definujme množiny $\mathcal{I}_0 = \mathcal{I} \cup \{0\}$ a $\mathcal{I}_0(\boldsymbol{x}) = \mathcal{I}(\boldsymbol{x}) \cup \{0\}$.

Nyní uveďme několik dodatečných předpokladů, které budou dále využívány pro reformulaci pravděpodobnostního omezení úlohy (3.2):

A1 Disperzní funkce $\boldsymbol{d}$ je $\mathcal{D}$-konvexní.

A1' Disperzní funkce $\boldsymbol{d}$ je $\mathcal{D}$-konvexní a pozitivně homogenní stupně 1.

A2 Množina $\boldsymbol{\Omega}$ je konvexní a uzavřená.

A2' Množina $\boldsymbol{\Omega}$ je konvexní a uzavřený kužel.

A3 Rodina pravděpodobnostních rozdělení $\mathcal{P}$ splňuje *Slaterovu podmínku* $\boldsymbol{\mu} \in \operatorname{int}\boldsymbol{\Omega}$ a $\boldsymbol{d}(\boldsymbol{\mu}) \prec_{\mathcal{D}} \boldsymbol{\sigma}$.

A4 Matice $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T}$ je konstantní $\forall \boldsymbol{x} \in \mathbb{R}^n$.

Pravděpodobnostní omezení z úlohy (3.2) je poměrně složité a reprezentuje nekonečně rozměrný problém. Pomocí následujících kroků lze přejít k ekvivalentnímu jednodušímu problému konečné dimenze:

- Infimum v omezení z úlohy (3.2) nahradíme konečně rozměrným problémem (P1) (Věta 3.3).
- Problém (P1) transformujeme do ekvivalentního minimalizačního konvexního problému (P2) (Věta 3.4).
- Z duality přejdeme od (P2) k maximalizačnímu problému (P3), jehož optimální hodnota musí být dle (3.2) alespoň $1-\epsilon$. Toto je ekvivalentní existenci přípustného řešení (P3) takovému, že účelová funkce (P3) je rovna alespoň $1-\epsilon$.

Tímto postupem lze nahradit složitou podmínku z úlohy (3.2) konečným počtem nerovností.

Věta 3.3 (Redukce do konečné dimenze, Hanasusanto a kol. (2017))**.** *Hodnota* $\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})]$ *z pravděpodobnostního omezení úlohy (3.2) je za podmínek A1, A2 a A3 rovna optimální hodnotě konečně rozměrného problému:*

$$\begin{aligned}
\min_{\lambda_i,\boldsymbol{\omega}_i} \quad & \lambda_0 \\
s.t. \quad & \lambda_i \in \mathbb{R}_+,\ \boldsymbol{\omega}_i \in \boldsymbol{\Omega},\ i \in \mathcal{I}_0(\boldsymbol{x}) \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i = 1 \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i \boldsymbol{\omega}_i = \boldsymbol{\mu} \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i \boldsymbol{d}(\boldsymbol{\omega}_i) \preceq_{\mathcal{D}} \boldsymbol{\sigma} \\
& \boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\omega}_i \geq u_i(\boldsymbol{x}), \forall i \in \mathcal{I}(\boldsymbol{x}).
\end{aligned} \tag{3.5}$$

Důkaz. Levá strana pravděpodobnostního omezení v úloze (3.2) je dle Theoremu 6 v článku Hanasusanto a kol. (2017) ekvivalentní problému:

$$\begin{aligned}
\max_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma}} \quad & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} - \boldsymbol{\sigma}^\top \boldsymbol{\gamma} \\
s.t. \quad & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^l,\ \boldsymbol{\gamma} \in \mathcal{D}^* \\
& \alpha + \boldsymbol{\omega}^\top \boldsymbol{\beta} - \boldsymbol{d}(\boldsymbol{\omega})^\top \boldsymbol{\gamma} \leq \mathbb{I}_{[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})]}, \forall \boldsymbol{\omega} \in \boldsymbol{\Omega}
\end{aligned}$$

Nerovnost obsahující na pravé straně indikátor množiny lze rozdělit na dva případy, což vede na úlohu:

$$\begin{aligned}
\max_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma}} \quad & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} - \boldsymbol{\sigma}^T \boldsymbol{\gamma} \\
s.t. \quad & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^l,\ \boldsymbol{\gamma} \in \mathcal{D}^* \\
& \alpha + \boldsymbol{\omega}_0^\top \boldsymbol{\beta} - \boldsymbol{d}(\boldsymbol{\omega}_0)^\top \boldsymbol{\gamma} \leq 1, \forall \boldsymbol{\omega}_0 \in \boldsymbol{\Omega}_0(\boldsymbol{x}) \\
& \alpha + \boldsymbol{\omega}_i^\top \boldsymbol{\beta} - \boldsymbol{d}(\boldsymbol{\omega}_i)^\top \boldsymbol{\gamma} \leq 0, \forall \boldsymbol{\omega}_i \in \boldsymbol{\Omega}_i(\boldsymbol{x}), i \in \mathcal{I}(\boldsymbol{x}),
\end{aligned}$$

kde

$$\boldsymbol{\Omega}_0(\boldsymbol{x}) = \boldsymbol{\Omega} \text{ a } \boldsymbol{\Omega}_i(\boldsymbol{x}) = \{\boldsymbol{\omega} \in \boldsymbol{\Omega} : \boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\omega} \geq u_i(\boldsymbol{x})\} \text{ pro } i \in \mathcal{I}(\boldsymbol{x}). \tag{3.6}$$

Výsledkem je optimalizační úloha s rozhodovacími proměnnými α, $\boldsymbol{\beta}$ a $\boldsymbol{\gamma}$ řešená za *pesimistických scénářů* $\boldsymbol{\omega}_i$, $i \in \mathcal{I}_0(\boldsymbol{x})$. Dle Theoremu 4.1 v článku Beck a Ben-Tal (2009) lze tento výstup přepsat do tvaru optimalizační úlohy:

$$\begin{aligned}
\min_{\lambda_i,\boldsymbol{\omega}_i} \quad & \lambda_0 \\
s.t. \quad & \lambda_i \in \mathbb{R}_+,\ \boldsymbol{\omega}_i \in \boldsymbol{\Omega}_i(\boldsymbol{x}),\ i \in \mathcal{I}_0(\boldsymbol{x}) \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i = 1 \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i \boldsymbol{\omega}_i = \boldsymbol{\mu} \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i \boldsymbol{d}(\boldsymbol{\omega}_i) \preceq_{\mathcal{D}} \boldsymbol{\sigma}.
\end{aligned}$$

Na závěr je omezení $\boldsymbol{\omega}_i \in \boldsymbol{\Omega}_i(\boldsymbol{x})$ nahrazeno jednodušším $\boldsymbol{\omega}_i \in \boldsymbol{\Omega}$ za přidání podmínky (viz definice množiny $\boldsymbol{\Omega}_i(\boldsymbol{x})$ (3.6)):

$$\boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\omega}_i \geq u_i(\boldsymbol{x}), \forall i \in \mathcal{I}(\boldsymbol{x}).$$

Celkem tedy bylo dokázáno, že hodnota $\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})]$ z pravděpodobnostního omezení v úloze (3.2) je rovna optimální hodnotě problému (3.5). □

Předmětem následující věty bude přeformulování úlohy (3.5) na konvexní problém. V této souvislosti definujme pro $\boldsymbol{d}(\boldsymbol{\omega})$ a $\lambda > 0$ funkci $\hat{\boldsymbol{d}}(\boldsymbol{\chi},\lambda)$ předpisem $\hat{\boldsymbol{d}}(\boldsymbol{\chi},\lambda) = \lambda \boldsymbol{d}(\frac{\chi}{\lambda})$ s úmluvou $\hat{\boldsymbol{d}}(\boldsymbol{\chi},0) = \lim_{\lambda\to 0^+} \lambda \boldsymbol{d}(\frac{\chi}{\lambda})$. Existence této limity je zaručena, pokud je funkce $\boldsymbol{d}(\boldsymbol{\omega})$ $\mathcal{D}$-konvexní a zdola polospojitá (viz Důsledek 8.5.2 v knize Rockafellar (1997)).

Věta 3.4 (Konvexní reformulace, Hanasusanto a kol. (2017))**.** *Redukovaná úloha (3.5) má za předpokladů A1, A2 a A3 stejnou optimální hodnotu jako následující konvexní problém:*

$$\begin{aligned}
\min_{\lambda_i,\boldsymbol{\chi}_i} \quad & \lambda_0 \\
s.t. \quad & \lambda_i \in \mathbb{R}_+,\ \boldsymbol{\chi}_i \in \mathbb{R}^l,\ i \in \mathcal{I}_0 \\
& \sum_{i\in\mathcal{I}_0} \lambda_i = 1 \\
& \frac{\boldsymbol{\chi}_i}{\lambda_i} \in \boldsymbol{\Omega}, \forall i \in \mathcal{I}_0 \\
& \sum_{i\in\mathcal{I}_0} \boldsymbol{\chi}_i = \boldsymbol{\mu} \\
& \sum_{i\in\mathcal{I}_0} \lambda_i \boldsymbol{d}(\frac{\boldsymbol{\chi}_i}{\lambda_i}) \preceq_{\mathcal{D}} \boldsymbol{\sigma} \\
& \boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\chi}_i \geq \lambda_i u_i(\boldsymbol{x}), \forall i \in \mathcal{I}
\end{aligned} \tag{3.7}$$

Důkaz. Pro každé přípustné řešení $(\lambda_i',\boldsymbol{\omega}_i')$ úlohy (3.5) je $(\lambda_i,\boldsymbol{\chi}_i) = (\lambda_i',\lambda_i'\boldsymbol{\omega}_i')$ pro $i \in \mathcal{I}_0(\boldsymbol{x})$ a $(\lambda_i,\boldsymbol{\chi}_i) = (0,\boldsymbol{0})$ pro $i \in \mathcal{I}_0\backslash\mathcal{I}_0(\boldsymbol{x})$ přípustné řešení úlohy (3.7) a dosahuje stejné hodnoty účelové funkce. Optimální hodnota úlohy (3.7) poskytuje dolní mez pro optimální hodnotu úlohy (3.5).

Dalším krokem je ukázat, že optimální hodnota úlohy (3.7) poskytuje také horní mez pro optimální hodnotu úlohy (3.5). K tomuto zadefinujme pomocný problém:

$$\begin{aligned}
\min_{\lambda_i,\boldsymbol{\chi}_i} \quad & \lambda_0 \\
s.t. \quad & \lambda_i \in \mathbb{R}_+,\ \boldsymbol{\chi}_i \in \mathbb{R}^l,\ i \in \mathcal{I}_0(\boldsymbol{x}) \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i = 1 \\
& \frac{\boldsymbol{\chi}_i}{\lambda_i} \in \boldsymbol{\Omega}, \forall i \in \mathcal{I}_0(\boldsymbol{x}) \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \boldsymbol{\chi}_i = \boldsymbol{\mu} \\
& \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i \boldsymbol{d}(\frac{\boldsymbol{\chi}_i}{\lambda_i}) \preceq_{\mathcal{D}} \boldsymbol{\sigma} \\
& \boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\chi}_i \geq \lambda_i u_i(\boldsymbol{x}), \forall i \in \mathcal{I}(\boldsymbol{x}),
\end{aligned} \tag{3.8}$$

kde jsou použity stejné indexové množiny $\mathcal{I}_0(\boldsymbol{x})$ a $\mathcal{I}(\boldsymbol{x})$ jako v problému (3.5). Nyní ukažme, že pro každé přípustné řešení úlohy (3.7) existuje přípustné řešení problému (3.8), které má stejnou hodnotu účelové funkce. Uvažujme tedy libovolné přípustné řešení $(\lambda'_i,\boldsymbol{\chi}'_i)$ problému (3.7). Poté definujme $(\lambda_i, \boldsymbol{\chi}_i)$ takto:

$$\lambda_0 = \lambda'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \lambda'_i,\ \lambda_i = \lambda'_i\ \forall i \in \mathcal{I}(\boldsymbol{x}),$$
$$\boldsymbol{\chi}_0 = \boldsymbol{\chi}'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \boldsymbol{\chi}'_i,\ \boldsymbol{\chi}_i = \boldsymbol{\chi}'_i\ \forall i \in \mathcal{I}(\boldsymbol{x})$$

a ukažme, že jsou přípustným řešením výše uvedeného problému. Pro $\lambda_i \geq 0$ a $i \in \mathcal{I}_0(\boldsymbol{x})$ dostáváme:

$$\sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i = \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda'_i = 1 \text{ a } \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \boldsymbol{\chi}_i = \sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \boldsymbol{\chi}'_i = \boldsymbol{\mu},$$

kde druhé rovnosti platí, neboť $(\lambda'_i,\boldsymbol{\chi}'_i)$ je přípustné řešení úlohy (3.7). Z $\mathcal{D}$-konvexity disperzní funkce $\boldsymbol{d}(\boldsymbol{\omega})$ vyplývá:

$$\sum_{i\in\mathcal{I}_0(\boldsymbol{x})} \lambda_i \boldsymbol{d}\left(\frac{\boldsymbol{\chi}_i}{\lambda_i}\right) = \left(\lambda'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \lambda'_i\right) \boldsymbol{d}\left(\frac{\boldsymbol{\chi}'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \boldsymbol{\chi}'_i}{\lambda'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \lambda'_i}\right) + \sum_{i\in\mathcal{I}(\boldsymbol{x})} \lambda'_i \boldsymbol{d}\left(\frac{\boldsymbol{\chi}'_i}{\lambda'_i}\right)$$
$$\preceq_{\mathcal{D}} \sum_{i\in\mathcal{I}_0} \lambda'_i \boldsymbol{d}\left(\frac{\boldsymbol{\chi}'_i}{\lambda'_i}\right) \preceq_{\mathcal{D}} \boldsymbol{\sigma}.$$

Dále víme, že $\frac{\boldsymbol{\chi}_i}{\lambda_i} = \frac{\boldsymbol{\chi}'_i}{\lambda'_i} \in \boldsymbol{\Omega}$ pro všechny $i \in \mathcal{I}(\boldsymbol{x})$. Také platí, že:

$$\begin{aligned}\frac{\boldsymbol{\chi}_0}{\lambda_0} &= \frac{\boldsymbol{\chi}'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \boldsymbol{\chi}'_i}{\lambda'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \lambda'_i} \\ &= \frac{\lambda'_0}{\lambda'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \lambda'_i}\frac{\boldsymbol{\chi}'_0}{\lambda'_0} + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \frac{\lambda'_i}{\lambda'_0 + \sum_{i\in\mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})} \lambda'_i}\frac{\boldsymbol{\chi}'_i}{\lambda'_i} \in \boldsymbol{\Omega},\end{aligned}$$

neboť $\boldsymbol{\Omega}$ je konvexní množina a $\frac{\boldsymbol{\chi}_0}{\lambda_0}$ lze vyjádřit jako konvexní kombinaci přípustných řešení úlohy (3.7). Splnění poslední podmínky přípustnosti je zcela zřejmé:

$$\boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\chi}_i = \boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\chi}'_i \geq \lambda'_i u_i(\boldsymbol{x}) = \lambda_i u_i(\boldsymbol{x})\ \forall i \in \mathcal{I}(\boldsymbol{x}).$$

Tudíž $(\lambda_i,\boldsymbol{\chi}_i)$ je přípustné řešení definované úlohy. Pro $(\lambda_i,\boldsymbol{\chi}_i)$ je hodnota účelové funkce rovna hodnotě účelové funkce pro hodnoty $(\lambda'_i,\boldsymbol{\chi}'_i)$ z úlohy (3.7), neboť platí, že $\lambda'_i = 0$ pro všechna $i \in \mathcal{I}\setminus\mathcal{I}(\boldsymbol{x})$. V opačném případě totiž $\boldsymbol{\omega}'_i = \frac{\boldsymbol{\chi}'_i}{\lambda'_i} \in \boldsymbol{\Omega}$ splňuje nerovnost $\boldsymbol{t}_i(\boldsymbol{x})^\top \boldsymbol{\omega}'_i \geq u_i(\boldsymbol{x})$, což je v rozporu s definicí indexové množiny $\mathcal{I}(\boldsymbol{x})$. To znamená, že řešení úlohy (3.7) dává horní mez pro optimální hodnotu úlohy (3.8).

Nyní ukažme, že řešení úlohy (3.8) vytváří horní mez pro optimální hodnotu úlohy (3.5). Předpokládejme, že $(\lambda_i,\boldsymbol{\chi}_i)$ je přípustné řešení úlohy (3.8). Na základě drobné modifikace Lemmatu 3.5 lze ukázat, že pro úlohu (3.8) existuje *Slaterův bod* $(\lambda'_i,\boldsymbol{\chi}'_i)$ s $\lambda'_i > 0\ \forall i \in \mathcal{I}_0(\boldsymbol{x})$. Pro konvexní úlohu (3.8) můžeme vytvořit konvexní kombinaci z bodů $(\lambda_i,\boldsymbol{\chi}_i)$ a $(\lambda'_i,\boldsymbol{\chi}'_i)$, která bude stále ležet v množině přípustných řešení a použijeme ji k vytvoření posloupnosti přípustných řešení:

$$(\lambda^k_i,\boldsymbol{\chi}^k_i) = (1 - \frac{1}{k})(\lambda_i,\boldsymbol{\chi}_i) + \frac{1}{k}(\lambda'_i,\boldsymbol{\chi}'_i),$$

jež konverguje k $(\lambda_i,\boldsymbol{\chi}_i)$ pro $k \to +\infty$ a splňuje $\lambda_i^k > 0\ \forall i \in \mathcal{I}_0(\boldsymbol{x})$. Odpovídající řešení $(\lambda_i^k,\boldsymbol{\omega}_i^k)$, kde $\boldsymbol{\omega}_i^k = \frac{\boldsymbol{\chi}_i^k}{\lambda_i^k}$ jsou přípustnými řešeními úlohy (3.5) a dosahují stejné hodnoty účelové funkce. To znamená, že optimální hodnota úlohy (3.8), a tedy i úlohy (3.7) tvoří horní odhad pro optimální hodnotu úlohy (3.5), což jsme chtěli dokázat.

Posledním krokem je ověření toho, zda tato přeformulovaná úloha (3.7) je skutečně konvexní. To znamená ověřit, zda je funkce $\hat{\boldsymbol{d}}(\boldsymbol{\chi},\lambda) = \lambda \boldsymbol{d}(\frac{\chi}{\lambda})$ $\mathcal{D}$-konvexní. Nechť $\lambda_1,\ \lambda_2 > 0$, $\boldsymbol{\chi}_1,\ \boldsymbol{\chi}_2 \in \mathbb{R}^l$ a $\theta \in [0,1]$, potom platí:

$$
\begin{aligned}
&\hat{\boldsymbol{d}}(\theta\boldsymbol{\chi}_1 + (1-\theta)\boldsymbol{\chi}_2,\theta\lambda_1 + (1-\theta)\lambda_2) \\
&= (\theta\lambda_1 + (1-\theta)\lambda_2)\boldsymbol{d}\left(\frac{\theta\boldsymbol{\chi}_1 + (1-\theta)\boldsymbol{\chi}_2}{\theta\lambda_1 + (1-\theta)\lambda_2}\right) \\
&= (\theta\lambda_1 + (1-\theta)\lambda_2)\mathbf{d}\left(\frac{\theta\lambda_1}{\theta\lambda_1 + (1-\theta)\lambda_2}\frac{\boldsymbol{\chi}_1}{\lambda_1} + \frac{(1-\theta)\lambda_2}{\theta\lambda_1 + (1-\theta)\lambda_2}\frac{\boldsymbol{\chi}_2}{\lambda_2}\right) \\
&\preceq_{\mathcal{D}} \theta\lambda_1\boldsymbol{d}\left(\frac{\boldsymbol{\chi}_1}{\lambda_1}\right) + (1-\theta)\lambda_2\boldsymbol{d}\left(\frac{\boldsymbol{\chi}_2}{\lambda_2}\right) = \theta\hat{\boldsymbol{d}}(\theta\boldsymbol{\chi}_1,\lambda_1) + (1-\theta)\hat{\boldsymbol{d}}(\boldsymbol{\chi}_2,\lambda_2),
\end{aligned}
$$

kde nerovnost plyne z $\mathcal{D}$-konvexity funkce $\boldsymbol{d}(\boldsymbol{\omega})$ (předpoklad A1). Funkce $\hat{\boldsymbol{d}}(\boldsymbol{\chi},\lambda)$ je tedy $\mathcal{D}$-konvexní. Úrovňové množiny součtu konvexních funkcí jsou konvexní, a tudíž množina přípustných řešení určená tímto omezením je konvexní. Omezení $\frac{\chi}{\lambda} \in \boldsymbol{\Omega}$ v úloze (3.7) vytváří konvexní množinu právě tehdy, pokud je množina $\boldsymbol{\Omega}$ konvexní, což zaručuje předpoklad A2. Zbylá omezení jsou lineární a tudíž je celá úloha konvexní. □

Zde ještě uveďme důkaz lemmatu, na které bylo odkazováno v průběhu důkazu předcházející věty.

Lemma 3.5. *Pro konvexní úlohu (3.7) existuje za splnění podmínek A1 a A3 Slaterův bod.*

Důkaz. Z definice indexové množiny $\mathcal{I}(\boldsymbol{x})$ (3.4) plyne, že $\forall i \in \mathcal{I}(\boldsymbol{x})\ \exists\boldsymbol{\omega}_i \in \boldsymbol{\Omega}$: $\boldsymbol{t}_i^\top(\boldsymbol{x})\boldsymbol{\omega}_i \geq u_i(\boldsymbol{x})$. Potom pro $i \in \mathcal{I}(\boldsymbol{x})$ a $\delta > 0$ malé volme $\boldsymbol{\chi}_i = \lambda_i\boldsymbol{\omega}_i$ a $\lambda_i = \delta$. Ostatní hodnoty jsou nastaveny jako $\lambda_i = 0$, $\boldsymbol{\chi}_i = 0$ pro $i \in \mathcal{I}\backslash\mathcal{I}(\boldsymbol{x})$ a $\lambda_0 = 1 - \delta\dim\mathcal{I}(\boldsymbol{x})$, $\boldsymbol{\chi}_0 = \boldsymbol{\mu} - \sum_{i\in\mathcal{I}(\boldsymbol{x})}\boldsymbol{\chi}_i$. Pro δ dostatečně malé platí, že $\frac{\chi_0}{\lambda_0} \in \operatorname{int}\boldsymbol{\Omega}$, neboť $\boldsymbol{\mu} \in \operatorname{int}\boldsymbol{\Omega}$, a $\sum_{i\in\mathcal{I}_0}\lambda_i\boldsymbol{d}(\frac{\chi_i}{\lambda_i}) \prec_{\mathcal{D}} \boldsymbol{\sigma}$, což plyne z $\boldsymbol{d}(\boldsymbol{\mu}) \prec_{\mathcal{D}} \boldsymbol{\sigma}$ a spojitosti disperzní funkce. Splnění ostatních omezujících podmínek je zřejmé:

$$
\begin{aligned}
\sum_{i\in\mathcal{I}_0}\lambda_i &= \sum_{i\in\mathcal{I}_0(\boldsymbol{x})}\lambda_i = 1 - \delta\dim\mathcal{I}(\boldsymbol{x}) + \sum_{i\in\mathcal{I}(\boldsymbol{x})}\delta = 1, \\
\sum_{i\in\mathcal{I}_0}\boldsymbol{\chi}_i &= \sum_{i\in\mathcal{I}_0(\boldsymbol{x})}\boldsymbol{\chi}_i = \boldsymbol{\mu} - \sum_{i\in\mathcal{I}(\boldsymbol{x})}\boldsymbol{\chi}_i + \sum_{i\in\mathcal{I}(\boldsymbol{x})}\boldsymbol{\chi}_i = \boldsymbol{\mu}.
\end{aligned}
$$

U afinních podmínek $\boldsymbol{t}_i(\boldsymbol{x})^\top\boldsymbol{\chi}_i \geq \lambda_i u_i(\boldsymbol{x})$ není ostrá nerovnost vyžadována (viz Definice 1.4). Tím je dokázáno, že $(\lambda_i,\boldsymbol{\chi}_i)$ je *Slaterovým bodem* úlohy (3.7). □

Obsahem další věty bude převedení konvexního problému (3.7) do kónického tvaru. Nejprve však uveďme krátké lemma, které bude použito v jejím důkazu.

Lemma 3.6. *Pro jakékoliv $\boldsymbol{\gamma} \in \mathcal{D}^*$ je duál konvexní funkce $\boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\omega})$ dán vztahem:*

$$(\boldsymbol{\gamma}^\top \boldsymbol{d})^*(\boldsymbol{\nu}) = \sigma_{\mathrm{epi}\, \boldsymbol{d}}(\boldsymbol{\nu}, -\boldsymbol{\gamma})$$

Důkaz. Pomocí definic konjugované a opěrné funkce lze dané výrazy vyjádřit jako:

$$\begin{aligned} (\boldsymbol{\gamma}^\top \boldsymbol{d})^*(\boldsymbol{\nu}) &= \sup_{\boldsymbol{\xi} \in \mathbb{R}^l} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\xi})\}, \\ \sigma_{\mathrm{epi}\, \boldsymbol{d}}(\boldsymbol{\nu}, -\boldsymbol{\gamma}) &= \sup_{\boldsymbol{\xi} \in \mathbb{R}^l, \boldsymbol{\eta} \in \mathbb{R}^d} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{\eta} : \boldsymbol{d}(\boldsymbol{\xi}) \preceq_{\mathcal{D}} \boldsymbol{\eta}\}. \end{aligned}$$

Rovnost mezi těmito dvěma výrazy budou dokázány postupně přes nerovnosti. První nerovnost plyne z toho, že supremum počítané přes větší množinu musí být nutně větší nebo rovno. Proto pro volbu $\boldsymbol{\eta} = \boldsymbol{d}(\boldsymbol{\xi})$ dostáváme:

$$\sup_{\boldsymbol{\xi} \in \mathbb{R}^l, \boldsymbol{\eta} \in \mathbb{R}^d} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{\eta} : \boldsymbol{d}(\boldsymbol{\xi}) \preceq_{\mathcal{D}} \boldsymbol{\eta}\} \geq \sup_{\boldsymbol{\xi} \in \mathbb{R}^l} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\xi})\}.$$

Pro důkaz druhé nerovnosti nejprve k výrazu uvnitř suprema přičteme a odečteme člen $\boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\xi})$:

$$\sup_{\boldsymbol{\xi} \in \mathbb{R}^l, \boldsymbol{\eta} \in \mathbb{R}^d} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\xi}) - \boldsymbol{\gamma}^\top (\boldsymbol{\eta} - \boldsymbol{d}(\boldsymbol{\xi})) : \boldsymbol{d}(\boldsymbol{\xi}) \preceq_{\mathcal{D}} \boldsymbol{\eta}\}.$$

Vzhledem k podmínce $\boldsymbol{d}(\boldsymbol{\xi}) \preceq_{\mathcal{D}} \boldsymbol{\eta}$ platí, že $\boldsymbol{\eta} - \boldsymbol{d}(\boldsymbol{\xi}) \in \mathcal{D}$. Jestliže $\boldsymbol{\gamma} \in \mathcal{D}^*$, tak musí platit, že $\gamma^\top (\boldsymbol{\eta} - \boldsymbol{d}(\boldsymbol{\xi})) \geq 0$. Pokud tedy bude tento člen ze suprema vypuštěn, tak se jeho hodnota nemůže snížit. To znamená, že:

$$\sup_{\boldsymbol{\xi} \in \mathbb{R}^l, \boldsymbol{\eta} \in \mathbb{R}^d} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{\eta} : \boldsymbol{d}(\boldsymbol{\xi}) \preceq_{\mathcal{D}} \boldsymbol{\eta}\} \leq \sup_{\boldsymbol{\xi} \in \mathbb{R}^l} \{\boldsymbol{\nu}^\top \boldsymbol{\xi} - \boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\xi})\}.$$

Tím je dokázána rovnost mezi oběma výrazy. □

Věta 3.7 (Kónická reformulace, Hanasusanto a kol. (2017))**.** *Za předpokladů A1', A2', A3 a A4 je pesimistické omezení z úlohy (3.2) splněno právě tehdy, pokud existují $\boldsymbol{\beta} \in \mathbb{R}^l$, $\boldsymbol{\gamma} \in \mathcal{D}^*$, $\tau_i \in \mathbb{R}_+$ pro $i \in \mathcal{I}$ a $\boldsymbol{\nu}_i \in \mathbb{R}^l$ pro $i \in \mathcal{I}_0$ splňující:*

$$\begin{aligned} 1 + \boldsymbol{\mu}^\top \boldsymbol{\beta} - \boldsymbol{\sigma}^\top \boldsymbol{\gamma} &\geq 1 - \epsilon \\ \boldsymbol{\nu}_0 - \boldsymbol{\beta} &\in \boldsymbol{\Omega}^* \\ \boldsymbol{\nu}_i - \boldsymbol{\beta} - \tau_i \boldsymbol{t}_i &\in \boldsymbol{\Omega}^*, \forall i \in \mathcal{I} \\ (-\boldsymbol{\nu}_i, \boldsymbol{\gamma}) &\in \mathrm{epi}^*\, \boldsymbol{d}, \forall i \in \mathcal{I}_0 \\ \left\| \begin{pmatrix} 2 \\ \tau_i - u_i(\boldsymbol{x}) \end{pmatrix} \right\|_2 &\leq \tau_i + u_i(\boldsymbol{x}), \forall i \in \mathcal{I}. \end{aligned} \tag{3.9}$$

Důkaz. Za platnosti předpokladů A1', A2' a A4 lze (3.7) zjednodušit do tvaru:

$$\begin{aligned} \min_{\lambda_i, \boldsymbol{\chi}_i} \quad & \lambda_0 \\ s.t. \quad & \lambda_i \in \mathbb{R}_+,\ \boldsymbol{\chi}_i \in \boldsymbol{\Omega},\ i \in \mathcal{I}_0 \\ & \sum_{i \in \mathcal{I}_0} \lambda_i = 1 \\ & \sum_{i \in \mathcal{I}_0} \boldsymbol{\chi}_i = \boldsymbol{\mu} \\ & \sum_{i \in \mathcal{I}_0} \boldsymbol{d}(\boldsymbol{\chi}_i) \preceq_{\mathcal{D}} \boldsymbol{\sigma} \\ & \boldsymbol{t}_i^\top \boldsymbol{\chi}_i \geq \lambda_i u_i(\boldsymbol{x}), \forall i \in \mathcal{I}, \end{aligned} \tag{3.10}$$

neboť dle předpokladu A1' platí, že $\lambda_i \boldsymbol{d}(\frac{\boldsymbol{\chi}_i}{\lambda_i}) = \boldsymbol{d}(\boldsymbol{\chi}_i)$. Na základě A2' víme, že množina $\boldsymbol{\Omega}$ je kužel, a tak lze podmínku $\frac{\boldsymbol{\chi}_i}{\lambda_i} \in \boldsymbol{\Omega}$ nahradit omezením $\boldsymbol{\chi}_i \in \boldsymbol{\Omega}$. Poslední předpoklad A4 znamená, že $\boldsymbol{t}_i(\boldsymbol{x}) = \boldsymbol{t}_i$. *Lagrangeova funkce* úlohy (3.10) má pak tvar:

$$\begin{aligned} L(\boldsymbol{\chi}_i,\lambda_i,\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau_i,i\in\mathcal{I}_0) &= \lambda_0 + \alpha(1-\sum_{i\in\mathcal{I}_0}\lambda_i) + \boldsymbol{\beta}^\top(\boldsymbol{\mu} - \sum_{i\in\mathcal{I}_0}\boldsymbol{\chi}_i) \\ &- \boldsymbol{\gamma}^\top(\boldsymbol{\sigma} - \sum_{i\in\mathcal{I}_0}\boldsymbol{d}(\boldsymbol{\chi}_i)) + \sum_{i\in\mathcal{I}}\tau_i(\lambda_i u_i(\mathbf{x}) - \mathbf{t}_i^\top\boldsymbol{\chi}_i) \\ &= \alpha + \boldsymbol{\mu}^\top\boldsymbol{\beta} - \boldsymbol{\sigma}^\top\boldsymbol{\gamma} + \lambda_0(1-\alpha) + \sum_{i\in\mathcal{I}}\lambda_i(\tau_i u_i(\boldsymbol{x}) - \alpha) \\ &+ (\boldsymbol{\gamma}^\top\boldsymbol{d}(\boldsymbol{\chi}_0) - \boldsymbol{\beta}^\top\boldsymbol{\chi}_0) + \sum_{i\in\mathcal{I}}(\boldsymbol{\gamma}^\top\boldsymbol{d}(\boldsymbol{\chi}_i) - \boldsymbol{\beta}^\top\boldsymbol{\chi}_i - \tau_i\mathbf{t}_i^\top\boldsymbol{\chi}_i) \end{aligned}$$

s multiplikátory $\alpha \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^l$, $\boldsymbol{\gamma} \in \mathcal{D}^*$, $\tau_i \in \mathbb{R}_+$, $i \in \mathcal{I}$. Pokud existuje *Slaterův bod* úlohy (3.10), pak je její optimální hodnota rovna:

$$\sup_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau_i,i\in\mathcal{I}} \inf_{\lambda_i\in\mathbb{R}_+,\boldsymbol{\chi}_i,i\in\mathcal{I}_0} L(\boldsymbol{\chi}_i,\lambda_i,\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau_i). \tag{3.11}$$

Důkaz toho, že pro úlohu (3.10) skutečně existuje *Slaterův bod* je analogií důkazu Lemmatu 3.5 s drobnými obměnami, proto ho zde nebudeme znovu provádět. Jelikož se jedná o konvexní úlohu, lze supremum ve vyjádření (3.11) nahradit maximem. Dále vyjděme ze vztahu:

$$\begin{aligned} \inf_{\lambda_i,\boldsymbol{\chi}_i,i\in\mathcal{I}_0} L(\boldsymbol{\chi}_i,\lambda_i,\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\tau_i) &= \alpha + \boldsymbol{\mu}^\top\boldsymbol{\beta} - \boldsymbol{\sigma}^\top\boldsymbol{\gamma} + \inf_{\lambda_0\in\mathbb{R}_+}\lambda_0(1-\alpha) \\ &+ \sum_{i\in\mathcal{I}}\inf_{\lambda_i\in\mathbb{R}_+}\lambda_i(\tau_i u_i(\boldsymbol{x}) - \alpha) + \inf_{\boldsymbol{\chi}_0\in\boldsymbol{\Omega}}(\boldsymbol{\gamma}^\top\boldsymbol{d}(\boldsymbol{\chi}_0) - \boldsymbol{\beta}^\top\boldsymbol{\chi}_0) \\ &+ \sum_{i\in\mathcal{I}}\inf_{\boldsymbol{\chi}_i\in\boldsymbol{\Omega}}(\boldsymbol{\gamma}^\top\boldsymbol{d}(\boldsymbol{\chi}_i) - \boldsymbol{\beta}^\top\boldsymbol{\chi}_i - \tau_i\boldsymbol{t}_i^\top\boldsymbol{\chi}_i), \end{aligned}$$

který bude maximalizován přes zbývající proměnné. Proto musí proměnná $\alpha \in \mathbb{R}$ splňovat omezení $\alpha \leq 1$ a $\alpha \leq \tau_i u_i(\mathbf{x})\ \forall i \in \mathcal{I}$. V opačném případě by členy obsahující tuto proměnou divergovaly do $-\infty$, což je vzhledem k následné maximalizaci nezajímavá varianta. Za platnosti vyřčených podmínek budou členy $\inf_{\lambda_0\in\mathbb{R}_+}\lambda_0(1-\alpha)$ a $\sum_{i\in\mathcal{I}}\inf_{\lambda_i\in\mathbb{R}_+}\lambda_i(\tau_i u_i(\boldsymbol{x}) - \alpha)$ nulové. Předposlední výraz lze upravit do tvaru:

$$\begin{aligned} \inf_{\boldsymbol{\chi}_0\in\boldsymbol{\Omega}}(\boldsymbol{\gamma}^\top\boldsymbol{d}(\boldsymbol{\chi}_0) - \boldsymbol{\beta}^\top\boldsymbol{\chi}_0) &= -\sup_{\boldsymbol{\chi}_0\in\mathbb{R}^l}(\boldsymbol{\beta}^\top\boldsymbol{\chi}_0 - \boldsymbol{\gamma}^\top\boldsymbol{d}(\boldsymbol{\chi}_0) - \delta_{\boldsymbol{\Omega}}(\boldsymbol{\chi}_0)) \\ &= -(\boldsymbol{\gamma}^\top\boldsymbol{d} + \delta_{\boldsymbol{\Omega}})^*(\boldsymbol{\beta}) \\ &= -\inf_{\boldsymbol{\nu}_0\in\mathbb{R}^l}((\boldsymbol{\gamma}^\top\boldsymbol{d})^*(\boldsymbol{\nu}_0) + \delta^*_{\boldsymbol{\Omega}}(\boldsymbol{\beta} - \boldsymbol{\nu}_0)) \\ &= \sup_{\boldsymbol{\nu}_0\in\mathbb{R}^l}(-\sigma_{\operatorname{epi}\boldsymbol{d}}(\boldsymbol{\nu}_0, -\boldsymbol{\gamma}) - \sigma_{\boldsymbol{\Omega}}(\boldsymbol{\beta} - \boldsymbol{\nu}_0)), \end{aligned}$$

kde v prvním kroku byla zaměněna množina, přes kterou je počítáno supremum, a přidán člen obsahující indikátor původní množiny. Další rovnost vychází z definice konjugované funkce. Konjugovaný součet dvou funkcí je vypočítán dle Theoremu

11.23(a) uvedeného v Rockafellar a Wets (2009) jako inf-konvoluce. V posledním kroku byl využit vztah mezi konjugovanou *indikátorovou funkcí* a *opěrnou funkcí* (viz vztah (1.3)) a Lemma 3.6. Analogicky platí pro poslední člen a $\forall i \in \mathcal{I}$:

$$\begin{aligned}
\inf_{\boldsymbol{\chi}_i \in \boldsymbol{\Omega}} (\boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\chi}_i) - \boldsymbol{\beta}^\top \boldsymbol{\chi}_i - \tau_i \boldsymbol{t}_i^\top \boldsymbol{\chi}_i) &= -\sup_{\boldsymbol{\chi}_i \in \mathbb{R}^l} (\boldsymbol{\beta}^\top \boldsymbol{\chi}_i + \tau_i \boldsymbol{t}_i^\top \boldsymbol{\chi}_i - \boldsymbol{\gamma}^\top \boldsymbol{d}(\boldsymbol{\chi}_i) - \delta_{\boldsymbol{\Omega}}(\boldsymbol{\chi}_i)) \\
&= -(\boldsymbol{\gamma}^\top \boldsymbol{d} + \delta_{\boldsymbol{\Omega}})^*(\boldsymbol{\beta} + \tau_i \boldsymbol{t}_i) \\
&= -\inf_{\boldsymbol{\nu}_i \in \mathbb{R}^l} ((\boldsymbol{\gamma}^\top \boldsymbol{d})^*(\boldsymbol{\nu}_i) + \delta^*_{\boldsymbol{\Omega}}(\boldsymbol{\beta} + \tau_i \boldsymbol{t}_i - \boldsymbol{\nu}_i)) \\
&= \sup_{\boldsymbol{\nu}_i \in \mathbb{R}^l} (-\sigma_{\operatorname{epi} \boldsymbol{d}}(\boldsymbol{\nu}_i, -\boldsymbol{\gamma}) - \sigma_{\boldsymbol{\Omega}}(\boldsymbol{\beta} + \tau_i \boldsymbol{t}_i - \boldsymbol{\nu}_i)).
\end{aligned}$$

Maximalizační úloha (3.11) má tedy tvar:

$$\begin{aligned}
\max_{\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_i, \boldsymbol{\nu}_i} \quad & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} - \boldsymbol{\sigma}^\top \boldsymbol{\gamma} - \sum_{i \in \mathcal{I}_0} \sigma_{\operatorname{epi} \boldsymbol{d}}(\boldsymbol{\nu}_i, -\boldsymbol{\gamma}) \\
& -\sigma_{\boldsymbol{\Omega}}(\boldsymbol{\beta} - \boldsymbol{\nu}_0) - \sum_{i \in \mathcal{I}} \sigma_{\boldsymbol{\Omega}}(\boldsymbol{\beta} + \tau_i \boldsymbol{t}_i - \boldsymbol{\nu}_i) \\
s.t. \quad & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^l,\ \boldsymbol{\gamma} \in \mathcal{D}^*,\ \tau_i \in \mathbb{R}_+,\ i \in \mathcal{I},\ \boldsymbol{\nu}_i \in \mathbb{R}^l,\ i \in \mathcal{I}_0 \\
& \alpha \leq 1 \\
& \alpha \leq \tau_i u_i(\boldsymbol{x}),\ \forall i \in \mathcal{I}.
\end{aligned}$$

Dle Lemmatu 1.11 je opěrná funkce uzavřeného konvexního kuželu $\mathcal{K}$ totožná s *indikátorovou funkcí* negativního duálního (polárního) kuželu $-\mathcal{K}^*$. Jelikož jsou množiny $\boldsymbol{\Omega}$ a $\operatorname{epi} \boldsymbol{d}$ na základě předpokladů A1' a A2' uzavřenými konvexními kuželi, lze psát $\sigma_{\boldsymbol{\Omega}}(\boldsymbol{\nu}) = \delta_{\boldsymbol{\Omega}^*}(-\boldsymbol{\nu})$ a $\sigma_{\operatorname{epi} \boldsymbol{d}}(\boldsymbol{\nu}) = \delta_{\operatorname{epi}^* \boldsymbol{d}}(-\boldsymbol{\nu})$. Na základě těchto souvislostí je možné duální úlohu přepsat do tvaru:

$$\begin{aligned}
\max_{\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_i, \boldsymbol{\nu}_i} \quad & \alpha + \boldsymbol{\mu}^\top \boldsymbol{\beta} - \boldsymbol{\sigma}^\top \boldsymbol{\gamma} \\
s.t. \quad & \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^l,\ \boldsymbol{\gamma} \in \mathcal{D}^*,\ \tau_i \in \mathbb{R}_+,\ i \in \mathcal{I},\ \boldsymbol{\nu}_i \in \mathbb{R}^l,\ i \in \mathcal{I}_0 \\
& \boldsymbol{\nu}_0 - \boldsymbol{\beta} \in \boldsymbol{\Omega}^* \\
& \boldsymbol{\nu}_i - \boldsymbol{\beta} - \tau_i \boldsymbol{t}_i \in \boldsymbol{\Omega}^*,\ \forall i \in \mathcal{I} \\
& (-\boldsymbol{\nu}_i, \boldsymbol{\gamma}) \in \operatorname{epi}^* \boldsymbol{d},\ \forall i \in \mathcal{I}_0 \\
& \alpha \leq 1 \\
& \alpha \leq \tau_i u_i(\boldsymbol{x}), \forall i \in \mathcal{I}.
\end{aligned} \tag{3.12}$$

V dalším kroku budou využity podmínky komplementarity, které platí díky Větě 1.5. Konkrétně pro každý pár řešení primární (3.9) a duální (3.12) úlohy musí platit $\lambda_0(1-\alpha) = 0$. Jelikož optimální hodnoty těchto úloh, které jsou na základě *silné duality* totožné, souvisí s hodnotou výrazu $\inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})]$, musí být splnění podmínky komplementarity zajištěno volbou $\alpha = 1$. Druhá varianta $\lambda_0 = 0$ by vedla k řešení $\inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})] = 0$ (viz Věta 3.3), které vzhledem k podmínce $\inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})] \geq 1 - \epsilon$ pro $\epsilon \in [0,1)$ nemá smysl uvažovat. Volbou $\alpha = 1$ se sníží počet omezení v úloze (3.12) a poslední z nich se transformuje do tvaru $1 \leq \tau_i u_i(\boldsymbol{x})\ \forall i \in \mathcal{I}$.

Hlavním cílem celé této kapitoly bylo zjistit, za jakých podmínek je splněno pesimistické pravděpodobnostní omezení úlohy (3.2). Maximum úlohy (3.12) je větší nebo rovno $1 - \epsilon$ právě tehdy, když existuje přípustné řešení, pro které

má účelová funkce hodnotu alespoň $1-\epsilon$. Tedy hledáme řešení splňující tyto podmínky:

$$\begin{aligned}
&1+\boldsymbol{\mu}^\top\boldsymbol{\beta}-\boldsymbol{\sigma}^\top\boldsymbol{\gamma}\geq 1-\epsilon\\
&\alpha\in\mathbb{R},\,\boldsymbol{\beta}\in\mathbb{R}^l,\,\boldsymbol{\gamma}\in\mathcal{D}^*,\,\tau_i\in\mathbb{R}_+,\,i\in\mathcal{I},\,\boldsymbol{\nu}_i\in\mathbb{R}^l,\,i\in\mathcal{I}_0\\
&\boldsymbol{\nu}_0-\boldsymbol{\beta}\in\boldsymbol{\Omega}^*\\
&\boldsymbol{\nu}_i-\boldsymbol{\beta}-\tau_i\boldsymbol{t}_i\in\boldsymbol{\Omega}^*,\,\forall i\in\mathcal{I}\\
&(-\boldsymbol{\nu}_i,\boldsymbol{\gamma})\in\mathrm{epi}^*\,\boldsymbol{d},\,\forall i\in\mathcal{I}_0\\
&1\leq\tau_i u_i(\boldsymbol{x}),\,\forall i\in\mathcal{I}.
\end{aligned}$$

Nakonec ještě ukážeme, že poslední podmínka v úloze (3.9) je totožná s podmínkou $1\leq\tau_i u_i(\boldsymbol{x})$. Vyjděme z tvaru podmínky, která je uvedena v úloze (3.9). Tuto nerovnost nejprve umocníme, což lze udělat, neboť $\tau_i\in\mathbb{R}_+$ a $1\leq\tau_i u_i(\boldsymbol{x})$. Z čehož dohromady plyne, že $u_i(\boldsymbol{x})\geq 0$. Další úpravy této nerovnosti jsou již zřejmé:

$$\begin{aligned}
\sqrt{4+(\tau_i-u_i(\boldsymbol{x}))^2} &\leq \tau_i+u_i(\boldsymbol{x})\\
4+\tau_i^2-2\tau_i u_i(\boldsymbol{x})+u_i^2(\boldsymbol{x}) &\leq \tau_i^2+2\tau_i u_i(\boldsymbol{x})+u_i^2(\boldsymbol{x})\\
1 &\leq \tau_i u_i(\boldsymbol{x}).
\end{aligned}$$

Tímto výsledkem je dokázáno, že pravděpodobnostní podmínka z úlohy (3.2) je ekvivalentní omezením (3.9). □

Celkem bylo pomocí Vět 3.3, 3.4 a 3.7 ukázáno, že úlohu (3.2) lze za platnosti požadovaných předpokladů upravit do ekvivalentní podoby:

$$\begin{aligned}
\min_{\boldsymbol{x}}\quad & f(\boldsymbol{x})\\
s.t.\quad & \boldsymbol{x}\in\boldsymbol{\mathcal{X}}\subseteq\mathbb{R}^n,\,\boldsymbol{\beta}\in\mathbb{R}^l,\,\boldsymbol{\gamma}\in\mathcal{D}^*,\\
& \tau_i\in\mathbb{R}_+,\,i\in\mathcal{I},\,\boldsymbol{\nu}_i\in\mathbb{R}^l,\,i\in\mathcal{I}_0\\
& 1+\boldsymbol{\mu}^\top\boldsymbol{\beta}-\boldsymbol{\sigma}^\top\boldsymbol{\gamma}\geq 1-\epsilon\\
& \boldsymbol{\nu}_0-\boldsymbol{\beta}\in\boldsymbol{\Omega}^*\\
& \boldsymbol{\nu}_i-\boldsymbol{\beta}-\tau_i\boldsymbol{t}_i\in\boldsymbol{\Omega}^*,\,\forall i\in\mathcal{I}\\
& (-\boldsymbol{\nu}_i,\boldsymbol{\gamma})\in\mathrm{epi}^*\,\boldsymbol{d},\,\forall i\in\mathcal{I}_0\\
& 1\leq\tau_i u_i(\boldsymbol{x}),\,\forall i\in\mathcal{I}.
\end{aligned}\tag{3.13}$$

Poslední omezení lze napsat i jako:

$$\left\|\begin{pmatrix}2\\ \tau_i-u_i(\boldsymbol{x})\end{pmatrix}\right\|_2\leq\tau_i+u_i(\boldsymbol{x}),$$

což je tvar tzv. *Lorentzova kužele*. Jedná se o typ kužele generovaného l^p-normami pro volbu $p=2$. Pro $p\geq 1$ je l^p-norma vektoru $\boldsymbol{x}\in\mathbb{R}^n$ definována jako:

$$\|\boldsymbol{x}\|_p=\sqrt[p]{\sum_{i=1}^n|x_i|^p}$$

s úmluvou $\|\boldsymbol{x}\|_\infty=\max_{i\in\{1,\dots,n\}}|x_i|$. Obecný kužel generovaný l^p-normou má pak tvar:

$$\mathcal{K}_p^{n+1}=\{(\boldsymbol{x},a)\in\mathbb{R}^{n+1}:\|\boldsymbol{x}\|_p\leq a\}.$$

Jedná se vlastně *epigraf* dané normy. Tyto kužele jsou konvexní, což plyne z trojúhelníkové nerovnosti pro normy a lze to lehce dokázat. Vezměme $(\boldsymbol{x},a)$, $(\boldsymbol{y},b) \in \mathcal{K}_p^{n+1}$ a $\lambda \in [0,1]$, potom platí, že:

$$\|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|_p \leq \lambda \|\boldsymbol{x}\|_p + (1-\lambda) \|\boldsymbol{y}\|_p \leq \lambda a + (1-\lambda)b,$$

kde první nerovnost plyne z trojúhelníkové nerovnosti a druhá z předpokladu $(\boldsymbol{x},a)$, $(\boldsymbol{y},b) \in \mathcal{K}_p^{n+1}$. Výsledkem je, že $(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda a + (1-\lambda)b) \in \mathcal{K}_p^{n+1}$, a tudíž jsou tyto kužele konvexní.

Další otázkou je, za jakých podmínek bude úloha (3.13) lineární konická. K tomu stačí požadovat, aby účelová funkce $f(\boldsymbol{x})$ byla afinní, neboť všechna omezení jsou v lineárním tvaru a množiny $\boldsymbol{\Omega}^*$, epi* $\boldsymbol{d}$ jsou kužele, což plyne z předpokladů A1', A2', Lemmatu 1.7 a faktu, že *epigraf* $\mathcal{D}$-konvexní pozitivně homogenní funkce (stupně 1) je kužel. Poslední omezení lze vyjádřit ve tvaru *Lorentzova kužele* a vzhledem k afinnímu tvaru funkcí $u_i(\boldsymbol{x})$ také linearita problému zůstává neporušena za předpokladu, že je množina $\boldsymbol{\mathcal{X}}$ polyedrická.

V případě ztráty linearity problému přichází na řadu otázka konvexity úlohy (3.13). V takovém případě musí být jak účelová funkce $f(\boldsymbol{x})$, tak množina $\mathcal{X}$ konvexní. Množiny $\boldsymbol{\Omega}^*$ a epi$(\boldsymbol{d})^*$ jsou za platnosti předpokladů A1' a A2' konvexní (viz Lemma 1.7). Jelikož $u_i(\boldsymbol{x}) > 0$, protože jinak by pro $\tau_i \geq 0$ nebylo možné splnit omezení $1 \leq \tau_i u_i(\boldsymbol{x})$, lze dané omezení přepsat do tvaru $\frac{1}{u_i(\boldsymbol{x})} \leq \tau_i$. Jedná se o *epigraf* složené funkce, kde vnější funkce $(\frac{1}{y})$ je hyperbola, která je konvexní pro $y > 0$, a vnitřní funkce $u_i(\boldsymbol{x}) > 0$ je dle předpokladů afinní. Složená funkce je tedy za daných podmínek konvexní a její *epigraf* je konvexní množina.

3.2 Ověření předpokladů

Při praktickém řešení optimalizačních úloh je důležitá jejich formulace tak, aby co nejlépe popisovaly daný problém a současně byly řešitelné a splňovaly předpoklady použitých metod. Otázkou, která je zcela zásadní při řešení úlohy (3.2), je výběr disperzní funkce $\boldsymbol{d}$. Jednou z možných voleb jsou tzv. *rizikové míry* (definované obvykle pro jednorozměrné náhodné veličiny $Z \in \mathbb{R}$), mezi které patří například variance, směrodatná odchylka, hodnota v riziku ($\mathrm{VaR}_\alpha(Z) = \inf\{z \in \mathbb{R} : F_Z(z) > \alpha\} = F_Z^{-1}(\alpha)$) nebo podmíněná hodnota v riziku ($\mathrm{CVaR}_\alpha(Z) = \mathbb{E}[Z|Z > \mathrm{VaR}_\alpha(Z)]$). K přechodu od úlohy (3.2) k úloze (3.13) je potřeba, aby vybraná riziková míra splňovala požadavek A1'. Z tohoto důvodu jsou zajímavé tzv. *koherentní míry rizika* $\rho(Z)$, které dle definice pro každé náhodné veličiny L, L_1, $L_2 \in \mathbb{R}$ a konstanty $a \geq 0$, $b \in \mathbb{R}$ splňují následující vztahy:

- monotonie: $L_1 \geq L_2$ *skoro jistě*: $\rho(L_1) \geq \rho(L_2)$,
- subaditivita: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$,
- pozitivní homogenita: $\rho(aL) = a\rho(L)$ a
- translační ekvivalence: $\rho(L + b) = \rho(L) + b$.

Koherentní míry rizika jsou pozitivně homogenní (stupně 1), což plyne přímo z definice, a konvexita je důsledkem subaditivity a pozitivní homogenity. Dalším požadavkem, na který je potřeba brát při výběru disperzní funkce zřetel, je snadný

výpočet duálu jejího *epigrafu* (epi* $\boldsymbol{d}$), což může být pro *koherentní míry rizika* celkem složitý úkol. O poznání jednodušší je výpočet duálu k *epigrafu* l^p-norem. Rizikové míry od nich odvozené ($\rho(Z) = \mathbb{E}\|Z\|_p$) sice nejsou koherentními, neboť nesplňují požadavek translační invariance:

$$\rho(Z+b) = \mathbb{E}\|Z+b\|_p \leq \mathbb{E}(\|Z\|_p + \|b\|_p) = \mathbb{E}\|Z\|_p + \mathbb{E}\|b\|_p = \rho(Z) + |b|,$$

neboť tato nerovnost skoro nikdy nebude rovností. Splňují však požadavek A1'. V souvislosti s výpočtem duálu k *epigrafu* l^p-normy je vhodné zavést duální normu vektoru $\boldsymbol{x} \in \mathbb{R}^n$ $\|\boldsymbol{x}\|_*$ předpisem:

$$\|\boldsymbol{x}\|_* = \sup_{\boldsymbol{y}\in\mathbb{R}^n} \{\boldsymbol{x}^\top \boldsymbol{y} : \|\boldsymbol{y}\| \leq 1\}. \tag{3.14}$$

Pomocí následujících dvou lemmat bude ukázáno, jak lze vypočítat duál k *epigrafu* l^p-normy. První lemma se zabývá vztahem mezi l^p-normami a jejich duálními podobami.

Lemma 3.8. *Pro* $\boldsymbol{x} \in \mathbb{R}^n$ *a* p, $q \geq 1$ *splňující* $\frac{1}{p} + \frac{1}{q} = 1$ *(s úmluvou* $\frac{1}{+\infty} = 0$*) platí* $\|\boldsymbol{x}\|_{p*} = \|\boldsymbol{x}\|_q$.

Důkaz. V prvním kroku se omezme na p, $q \in (1, +\infty)$, které jsou navíc *Hölderovsky konjugované* ($\frac{1}{p} + \frac{1}{q} = 1$). Dále lze bez újmy na obecnosti předpokládat, že $\boldsymbol{x} \neq \boldsymbol{0}$. V opačném případě by vztah mezi normami platil triviálně. Začneme nalezením horní meze pro $\|\boldsymbol{x}\|_{p*}$:

$$\boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^n x_i y_i \leq \sum_{i=1}^n |x_i y_i| \leq \|\boldsymbol{x}\|_q \|\boldsymbol{y}\|_p \leq \|\boldsymbol{x}\|_q .$$

Druhá nerovnost plyne z *Hölderovy nerovnosti* a třetí z omezení $\|\boldsymbol{y}\|_p \leq 1$ (viz vztah (3.14)).

K dokázání rovnosti stačí najít vektor $\boldsymbol{y} : \|\boldsymbol{y}\|_p \leq 1$, pro který platí, že $\boldsymbol{x}^\top \boldsymbol{y} = \|\boldsymbol{x}\|_q$. Nejprve definujme pomocný vektor $\boldsymbol{z}$ předpisem $z_i = \operatorname{sgn}(x_i)|x_i|^{q-1}$ $\forall i \in \{1,...,n\}$. Potom lze s využitím vztahu $\operatorname{sgn}(x_i)x_i = |x_i|$ psát:

$$\boldsymbol{x}^\top \boldsymbol{z} = \sum_{i=1}^n x_i z_i = \sum_{i=1}^n x_i \operatorname{sgn}(x_i)|x_i|^{q-1} = \sum_{i=1}^n |x_i|^q = \|\boldsymbol{x}\|_q^q . \tag{3.15}$$

Dále pro $\|\boldsymbol{z}\|_p^p$ platí:

$$\|\boldsymbol{z}\|_p^p = \sum_{i=1}^n |z_i|^p = \sum_{i=1}^n |\operatorname{sgn}(x_i)|x_i|^{q-1}|^p = \sum_{i=1}^n |x_i|^{p(q-1)} = \sum_{i=1}^n |x_i|^q = \|\boldsymbol{x}\|_q^q , \tag{3.16}$$

neboť pro *Hölderovsky konjugované* mocniny p a q je $p(q-1) = q$. Nyní vezměme vektor $\boldsymbol{y} = \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|_p}$. Tato definice je korektní, protože vzhledem k předpokladu $\boldsymbol{x} \neq \boldsymbol{0}$ je $\|\boldsymbol{z}\|_p > 0$. Za využití vztahů (3.15) a (3.16) dostáváme:

$$\boldsymbol{x}^\top \boldsymbol{y} = \frac{1}{\|\boldsymbol{z}\|_p} \boldsymbol{x}^\top \boldsymbol{z} = \frac{1}{\|\boldsymbol{x}\|_q^{\frac{q}{p}}} \|\boldsymbol{x}\|_q^q = \|\boldsymbol{x}\|_q^{q-\frac{q}{p}} = \|\boldsymbol{x}\|_q ,$$

protože pro p a q *Hölderovsky konjugované* platí, že $q - \frac{q}{p} = q(1 - \frac{1}{p}) = \frac{q}{q} = 1$. Tím byl nalezen vektor $\boldsymbol{y} : \|\boldsymbol{y}\|_p \leq 1$, který splňuje $\boldsymbol{x}^\top \boldsymbol{y} = \|\boldsymbol{x}\|_q$, a dokázáno, že $\|\boldsymbol{x}\|_{p*} = \|\boldsymbol{x}\|_q$ pro *Hölderovsky konjugované* p, $q \in (1, +\infty)$.

Nyní zbývá ještě dokončit důkaz pro normy $\|\boldsymbol{x}\|_1$ a $\|\boldsymbol{x}\|_\infty$. Začneme l^1-normou a výpočtem jejího duálu dle předpisu:

$$\|\boldsymbol{x}\|_{1*} = \sup_{\boldsymbol{y} \in \mathbb{R}^n} \{\boldsymbol{x}^\top \boldsymbol{y} : \|\boldsymbol{y}\|_1 \leq 1\}.$$

V tomto případě je dosazeno suprema pro volbu vektoru $\boldsymbol{y} : y_i = 0 \; \forall i \neq k$, kde k je index, pro který platí, že $|x_k| = \max_{i \in \{1,\ldots,n\}} |x_i|$, a $y_k = \operatorname{sgn}(x_k)$. Potom:

$$\|\boldsymbol{x}\|_{1*} = |x_k| = \max_{i \in \{1,\ldots,n\}} |x_i| = \|\boldsymbol{x}\|_\infty .$$

Pro l^∞-normu platí vztah:

$$\|\boldsymbol{x}\|_{\infty*} = \sup_{\boldsymbol{y} \in \mathbb{R}^n} \{\boldsymbol{x}^\top \boldsymbol{y} : \|\boldsymbol{y}\|_\infty \leq 1\}.$$

Nyní volme vektor $\boldsymbol{y}$, jež splňuje $y_i = \operatorname{sgn}(x_i) \; \forall i \in \{1,\ldots,n\}$ a pro který dosáhneme suprema:

$$\|\boldsymbol{x}\|_{\infty*} = \sum_{i \in \{1,\ldots,n\}} |x_i| = \|\boldsymbol{x}\|_1 .$$

Tím je lemma dokázáno pro p, $q \geq 1$. □

Následující lemma dává návod, jak určit duál k *epigrafu* l^p-normy.

Lemma 3.9. *Pro l^p-normy a p, $q \geq 1$ Hölderovsky konjugované platí:*

$$\operatorname{epi}^* \|\cdot\|_p = \operatorname{epi} \|\cdot\|_q .$$

Důkaz. Nejprve napišme definice obou množin:

$$\begin{aligned} \operatorname{epi}^* \|\cdot\|_p &= \{(\boldsymbol{x},a) \in \mathbb{R}^{n+1} : \boldsymbol{x}^\top \boldsymbol{y} + ab \geq 0 \; \forall (\boldsymbol{y},b) \in \mathbb{R}^{n+1} : \|\boldsymbol{y}\|_p \leq b\}, \\ \operatorname{epi} \|\cdot\|_q &= \{(\boldsymbol{x},a) \in \mathbb{R}^{n+1} : \|\boldsymbol{x}\|_q \leq a\}. \end{aligned}$$

Začněme s dokazováním inkluze $\operatorname{epi} \|\cdot\|_q \subseteq \operatorname{epi}^* \|\cdot\|_p$. Vezměme tedy $(\boldsymbol{x},a) \in \operatorname{epi} \|\cdot\|_q$ a $(\boldsymbol{y},b) \in \operatorname{epi} \|\cdot\|_p$, pro které platí $\|\boldsymbol{x}\|_q \leq a$ a $\|\boldsymbol{y}\|_p \leq b$. Všechny hodnoty v těchto dvou nerovnostech jsou nezáporné, a tudíž je lze dát dohromady:

$$ab \geq \|\boldsymbol{x}\|_q \|\boldsymbol{y}\|_p = \|-\boldsymbol{x}\|_q \|\boldsymbol{y}\|_p \geq -\boldsymbol{x}^\top \boldsymbol{y}.$$

Rovnost platí díky invariantnosti normy vůči změně znaménka vektoru a poslední nerovnost plyne z *Hölderovy nerovnosti*. Získaný vztah lze přepsat do tvaru:

$$\boldsymbol{x}^\top \boldsymbol{y} + ab \geq 0, \tag{3.17}$$

čímž je ověřeno, že $(\boldsymbol{x},a) \in \operatorname{epi}^* \|\cdot\|_p$, a dokázána první inkluze.

Pro dokázání druhé inkluze vezměme body $(\boldsymbol{x},a) \in \operatorname{epi}^* \|\cdot\|_p$ a $(\boldsymbol{y},b)$, kde $\boldsymbol{y} = \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|_p}$ pro $\boldsymbol{z} : z_i = -\operatorname{sgn}(x_i)|x_i|^{q-1} \; \forall i \in \{1,\ldots,n\}$ a $b = 1$. Protože $\|\boldsymbol{y}\|_p \leq b$,

tak pro dané body platí vztah (3.17). Pro *Hölderovsky konjugovanou* dvojici p a q lze psát:

$$\|\boldsymbol{z}\|_p^p = \sum_{i=1}^{n} |x_i|^{p(q-1)} = \|\boldsymbol{x}\|_q^q,$$

neboť $p(q-1) = q$. Dále lze skalární součin mezi vektory $\boldsymbol{x}$ a $\boldsymbol{y}$ vyjádřit jako:

$$\begin{aligned} \boldsymbol{x}^\top \boldsymbol{y} &= \frac{1}{\|\boldsymbol{z}\|_p} \sum_{i=1}^{n} x_i z_i = -\frac{1}{(\|\boldsymbol{z}\|_p^p)^{\frac{1}{p}}} \sum_{i=1}^{n} |x_i|^q \\ &= -\frac{\|\boldsymbol{x}\|_q^q}{\|\boldsymbol{x}\|_q^{\frac{q}{p}}} = -\|\boldsymbol{x}\|_q^{q-\frac{q}{p}} = -\|\boldsymbol{x}\|_q. \end{aligned} \tag{3.18}$$

Zkombinováním vztahů (3.17) a (3.18) dostaneme pro volbu $b = 1$:

$$\|\boldsymbol{x}\|_q = -\boldsymbol{x}^\top \boldsymbol{y} \leq ab = a.$$

Tudíž je bod $(\boldsymbol{x},a) \in \operatorname{epi}^* \|\cdot\|_p$ také prvkem množiny $\operatorname{epi} \|\cdot\|_q$. Tím je dokázána druhá inkluze, a tedy i rovnost mezi množinami.

Ještě zbývá dokončit důkaz pro $p = 1$ a $q = +\infty$. Začneme s inkluzí $\operatorname{epi} \|\cdot\|_\infty \subseteq \operatorname{epi}^* \|\cdot\|_1$. Vezměme $(\boldsymbol{x},a) \in \operatorname{epi} \|\cdot\|_\infty$ a $(\boldsymbol{y},b) \in \operatorname{epi} \|\cdot\|_1$. Vzhledem k nezápornosti norem a hodnot a, b platí:

$$\begin{aligned} ab &\geq \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1 = \max_{i\in\{1,\dots,n\}} |x_i| \sum_{i\in\{1,\dots,n\}} |y_i| \\ &\geq \sum_{i\in\{1,\dots,n\}} |x_i y_i| \geq -\sum_{i\in\{1,\dots,n\}} x_i y_i = -\boldsymbol{x}^\top \boldsymbol{y}. \end{aligned}$$

Z toho vyplývá, že $(\boldsymbol{x},a) \in \operatorname{epi}^* \|\cdot\|_1$, protože $\boldsymbol{x}^\top \boldsymbol{y} + ab \geq 0 \ \forall (\boldsymbol{y},b) \in \operatorname{epi} \|\cdot\|_1$.

Pro dokázání druhé inkluze ($\operatorname{epi}^* \|\cdot\|_1 \subseteq \operatorname{epi} \|\cdot\|_\infty$) volme $(\boldsymbol{y},1) \in \operatorname{epi} \|\cdot\|_1$ takový, že $y_i = 0 \ \forall i \neq k$, kde pro k platí, že $|x_k| = \max_{i\in\{1,\dots,n\}} |x_i|$ a $y_k = -\operatorname{sgn}(x_k)$. Potom pro $(\boldsymbol{x},a) \in \operatorname{epi}^* \|\cdot\|_1$ platí:

$$0 \leq \boldsymbol{x}^\top \boldsymbol{y} + a \cdot 1 = -\operatorname{sgn}(x_k) x_k + a = -\max_{i\in\{1,\dots,n\}} |x_i| + a = -\|\boldsymbol{x}\|_\infty + a,$$

tudíž je $(\boldsymbol{x},a) \in \operatorname{epi}^* \|\cdot\|_1$ také prvkem množiny $\operatorname{epi} \|\cdot\|_\infty$ ($\|\boldsymbol{x}\|_\infty \leq a$). Tím je dokázána rovnost mezi množinami. Platnost vztahu $\operatorname{epi}^* \|\cdot\|_\infty = \operatorname{epi} \|\cdot\|_1$ plyne z toho, že kužele generované l^p-normami jsou konvexní a uzavřené, a tudíž pro ně platí, že $(\operatorname{epi}^* \|\cdot\|_p)^* = \operatorname{epi} \|\cdot\|_p$. □

3.3 Optimistický scénář

V některých případech může být vhodnější místo *scénářů pesimistických* naopak uvažovat *scénáře optimistické*. Potom má dané pravděpodobnostní omezení tvar:

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{g}(\boldsymbol{x},\boldsymbol{\omega}) \geq 0] \geq 1-\epsilon. \tag{3.19}$$

Obdobně jako výše lze výraz na levé straně tohoto omezení se speciálním tvarem funkce $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{\omega}) = \boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega}$ v několika krocích přetransformovat do jednoduššího problému. Za platnosti předpokladů A1, A2 a A3 lze postupně přejít ke konečně rozměrnému konvexnímu problému. Pokud jsou použity striktnější předpoklady A1', A2' spolu s předpokladem A4, lze omezení (3.19) převést na ekvivalentní problém hledání $\boldsymbol{\omega}_i \in \boldsymbol{\Omega}$, $i \in \{0,1\}$ splňujících následující vztahy:

$$\begin{aligned} &\boldsymbol{T}\boldsymbol{\omega}_0 \leq (1-\epsilon)\boldsymbol{u}(\boldsymbol{x}) \\ &\sum_{i\in\{0,1\}} \boldsymbol{\omega}_i = \boldsymbol{\mu} \\ &\sum_{i\in\{0,1\}} \boldsymbol{d}(\boldsymbol{\omega}_i) \preceq_{\mathcal{D}} \boldsymbol{\sigma}. \end{aligned} \tag{3.20}$$

4. Aplikační část

V předešlých kapitolách práce byly připomenuty potřebné znalosti z *teorie optimalizace*, které byly následně aplikovány na úlohy s pravděpodobnostními omezeními, jež jsou hlavní náplní tohoto textu. Konkrétně se jedná o úlohy, ve kterých pravděpodobnostní rozdělení náhodných veličin definujících množinu přípustných řešení není přímo známo. Známá je pouze rodina pravděpodobnostních rozdělení, určena například předpisem (3.3). Jako možné přístupy řešení tohoto problému byly vybrány metody tzv. *pesimistických* a *optimistických scénářů*, kdy je příslušné pravděpodobnostní omezení formulováno ve tvaru (3.1), respektive (3.19). Dosavadním hlavním teoretickým výsledkem je možnost transformace takové optimalizační úlohy ve speciálním tvaru (3.2), za splnění jistých předpokladů, do výpočetně přijatelnější podoby (3.13). Pro případ *optimistického scénáře* lze pravděpodobnostní omezení přepsat do tvaru (3.20).

Optimalizační úlohy jsou spojeny s celou řadou oborů lidské činnosti, jako je logistika Cordeau a kol. (2007), finance Dert a Oldenkamp (2000), tvorba sítí Wang (2007), projektový management Shen a kol. (2010) a další. V této práci budou dosažené výsledky aplikovány na problémy z oblasti *optimalizace portfolia* a *zpracování obrazu*.

4.1 Optimalizace portfolia

V použitém modelu je portfolio chápáno jako soubor n aktiv držených jedním investorem. Zastoupení jednotlivých aktiv v portfoliu je dáno vektorem vah $\boldsymbol{x} \in \mathbb{R}^n$. Pokud nejsou na váhy x_i, $i \in \{1,...,n\}$ kladeny žádné dodatečné podmínky, tak předpokládáme, že tzv. *prodeje na krátko* jsou povoleny. V opačném případě je potřeba vektor vah omezit podmínkou $\boldsymbol{x} \in \mathbb{R}^n_+$. Množství aktiv v portfoliu je určeno bohatstvím investora w, které je ochoten do aktiv vložit. Tato skutečnost lze vyjádřit podmínkou $\mathbb{1}^\top \boldsymbol{x} = w$, kde $\mathbb{1}$ je vektor jedniček. Výnosy aktiv $\boldsymbol{\rho}$ jsou reprezentovány náhodným vektorem se střední hodnotou $\mathbb{E}\boldsymbol{\rho} = \boldsymbol{r}$ a kovariančí maticí $\operatorname{var}\boldsymbol{\rho} = \mathbb{V}$.

Optimalizace portfolia je proces, kdy se investor snaží alokovat své bohatství w mezi aktiva v portfoliu podle svých cílů, které jsou dány snahou dosáhnout jistého předem určeného výnosu r_e s minimálním možným rizikem, což je standardní chování popsané již v článku Markowitz (1952). Měřítkem rizika je rozptyl výnosu portfolia $\operatorname{var}\boldsymbol{\rho}^\top \boldsymbol{x} = \boldsymbol{x}^\top \mathbb{V} \boldsymbol{x}$. Místo náhodného vektoru $\boldsymbol{\rho}$ lze pracovat s centrovaným vektorem $\boldsymbol{\omega} = \boldsymbol{\rho} - \boldsymbol{r}$ (se stejnou kovariančí maticí $\mathbb{V}$), který vyjadřuje perturbace od střední hodnoty výnosů. Z těchto představ vychází *Markowitzův model portfolia*, jež odpovídá optimalizační úloze:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^\top \mathbb{V} \boldsymbol{x} \\ s.t. \quad & \boldsymbol{x} \in \mathbb{R}^n \\ & \mathbb{1}^\top \boldsymbol{x} = w \\ & \boldsymbol{r}^\top \boldsymbol{x} \geq r_e. \end{aligned} \tag{4.1}$$

Tento model předpokládá znalost parametrů ($\mathbb{E}\boldsymbol{\rho}$, $\operatorname{var}\boldsymbol{\rho}$) pravděpodobnostního rozdělení náhodného vektoru $\boldsymbol{\rho}$. Jedná se o konvexní problém, neboť všechna

omezení jsou v lineárním tvaru a kovarianční matice $\mathbb{V}$ je z definice pozitivně semidefinitní.

Nyní bude problém transformován do řeči *pesimistických scénářů*, kdy je do úlohy přidán náhodný efekt v podobě perturbací $\boldsymbol{\omega}$:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^\top \mathbb{V} \boldsymbol{x} \\ s.t. \quad & \boldsymbol{x} \in \mathbb{R}^n \\ & \mathbb{1}^\top \boldsymbol{x} = w \\ & \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{x}^\top (\boldsymbol{r} + \boldsymbol{\omega}) > r_e] \geq 1 - \epsilon, \end{aligned} \tag{4.2}$$

kde pro předem zvolenou disperzní funkci $\boldsymbol{d}(\boldsymbol{\omega}) : \mathbb{R}^n \to \mathbb{R}^d$ a parametr $\boldsymbol{\sigma} \in \mathbb{R}^d$ je rodina pravděpodobnostních rozdělení $\mathcal{P}$ dána jako:

$$\mathcal{P} = \{\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^n) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\omega}] = \mathbf{0},\, \mathbb{E}_{\mathbb{P}}[\boldsymbol{d}(\boldsymbol{\omega})] \leq \boldsymbol{\sigma}\}.$$

Pro převedení úlohy (4.2) do tvaru (3.13) je nutné identifikovat jednotlivé objekty pro tento konkrétní případ:

$$\begin{aligned} \boldsymbol{T}(\boldsymbol{x}) &= -\boldsymbol{x}^\top \\ \boldsymbol{u}(\boldsymbol{x}) &= \boldsymbol{x}^\top \boldsymbol{r} - r_e. \end{aligned}$$

Tím je dosaženo požadovaného tvaru pravděpodobnostního omezení úlohy (4.2):

$$\inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})] \geq 1 - \epsilon.$$

Při transformaci optimalizační úlohy ve tvaru *pesimistických scénářů* do podoby (3.13) byl ve Větě 3.7 použitý předpoklad A4 (konstantní matice $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T}$), který je v tomto případě porušen. Tento předpoklad však pouze zaručuje lineární tvar výsledné úlohy a nijak nesouvisí se samotným převodem úlohy do tvaru (3.13), který tedy může být učiněn i v tomto případě. Za disperzní funkci byla zvolena l^2-norma, která splňuje požadavek pozitivní homogenity (stupně 1), konvexitu a duál jejího *epigrafu* lze spočítat velice snadno, neboť platí $\operatorname{epi}^* \|\cdot\|_2 = \operatorname{epi} \|\cdot\|_2$ (viz Lemma 3.9). Kužel $\mathcal{D}$ je totožný se svým duálem a rovný $\mathbb{R}_+$. Konečně nosičem náhodného vektoru $\boldsymbol{\omega}$ je celý prostor $\mathbb{R}^n$, jehož duálem je bod $\mathbf{0} \in \mathbb{R}^n$ ($\boldsymbol{\Omega}^* = \{\mathbf{0}\}$). Horní hranice pro volbu parametru σ, který bude v tomto případě jednorozměrný, je tvořena stopou kovarianční matice $\mathbb{V}$ ($\sqrt{\operatorname{Tr} \mathbb{V}}$), neboť pro zvolenou disperzní funkci a centrovaný náhodný vektor $\boldsymbol{\omega}$ platí:

$$\mathbb{E} d(\boldsymbol{\omega}) = \mathbb{E} \|\boldsymbol{\omega}\|_2 \leq \sqrt{\mathbb{E} \sum_{i=1}^{n} \omega_i^2} = \sqrt{\sum_{i=1}^{n} \mathbb{E} \omega_i^2} = \sqrt{\operatorname{Tr} \mathbb{V}},$$

Při tomto výpočtu byla využita linearita střední hodnoty a *Jensenova nerovnost* pro konkávní funkci (druhá odmocnina). Nyní je vše připraveno k přepsání

úlohy (4.2) dle (3.13) do tvaru:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^\top \mathbb{V} \boldsymbol{x} \\ s.t. \quad & \boldsymbol{x} \in \mathbb{R}^n,\, \gamma \in \mathbb{R}_+, \\ & \tau \in \mathbb{R}_+,\, \boldsymbol{\nu}_0,\, \boldsymbol{\nu}_1, \in \mathbb{R}^n \\ & \mathbb{1}^\top \boldsymbol{x} = w \\ & 1 - \sigma\gamma \geq 1 - \epsilon \\ & \boldsymbol{\nu}_1 - \boldsymbol{\nu}_0 + \tau \boldsymbol{x} = \mathbf{0} \\ & \|\boldsymbol{\nu}_0\|_2 \leq \gamma \\ & \|\boldsymbol{\nu}_1\|_2 \leq \gamma \\ & 1 \leq \tau(\boldsymbol{x}^\top \boldsymbol{r} - r_e), \end{aligned} \tag{4.3}$$

kde byla ještě dimenze problému redukována eliminováním proměnné $\boldsymbol{\beta}$ na základě omezení $\boldsymbol{\nu}_0 - \boldsymbol{\beta} = \mathbf{0}$. Vzhledem k porušení předpokladu A4 již není daný problém konvexní, a tudíž není zaručeno, že nalezené minimum bude minimem globálním.

Za totožných úvah a požadavků lze k úloze přistoupit pomocí metody *optimistických scénářů* (viz vztah (3.20)), která vede na úlohu:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^\top \mathbb{V} \boldsymbol{x} \\ s.t. \quad & \boldsymbol{x} \in \mathbb{R}^n,\, \boldsymbol{\nu}_0,\, \boldsymbol{\nu}_1 \in \mathbb{R}^n \\ & \mathbb{1}^\top \boldsymbol{x} = w \\ & -\boldsymbol{x}^\top \boldsymbol{\nu}_0 \leq (1 - \epsilon)(\boldsymbol{x}^\top \boldsymbol{r} - r_e) \\ & \sum_{i \in \{0,1\}} \boldsymbol{\nu}_i = \boldsymbol{\mu} \\ & \sum_{i \in \{0,1\}} \|\boldsymbol{\nu}_i\|_2 \leq \sigma. \end{aligned}$$

S využitím volby $\boldsymbol{\mu} = \mathbf{0}$ a druhé rovnosti lze redukovat počet proměnných, neboť $\boldsymbol{\nu}_0 = -\boldsymbol{\nu}_1$. Výsledná úloha má tvar:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^\top \mathbb{V} \boldsymbol{x} \\ s.t. \quad & \boldsymbol{x} \in \mathbb{R}^n,\, \boldsymbol{\nu} \in \mathbb{R}^n \\ & \mathbb{1}^\top \boldsymbol{x} = w \\ & (1 - \epsilon) r_e \leq \boldsymbol{x}^\top ((1 - \epsilon)\boldsymbol{r} + \boldsymbol{\nu}) \\ & 2 \|\boldsymbol{\nu}\|_2 \leq \sigma. \end{aligned} \tag{4.4}$$

Ze stejného důvodu jako výše tento problém není konvexní.

Zajímavá situace nastává pro $\sigma = 0$. Následující argument je pouze formální, neboť pro tento případ není splněna *Slaterova podmínka* ($d(\boldsymbol{\mu}) < \sigma$), takže správný postup by byl uvažovat limitu $\sigma \to 0$. Nejprve začněme *optimistickou variantou* (4.4). Při volbě $\sigma = 0$ z poslední nerovnosti plyne, že $\|\boldsymbol{\nu}\|_2 \leq 0$. To znamená, že $\boldsymbol{\nu} = \mathbf{0}$. Po dosazení tohoto vztahu do předchozí nerovnosti a zkrácení kladného výrazu $1 - \epsilon$ na obou stranách, vznikne podmínka $r_e \leq \boldsymbol{x}^\top \boldsymbol{r}$, která je stejná jako v případě klasické úlohy (4.1). U *pesimistické varianty* (4.3) při volbě $\sigma = 0$ vypadne z první nerovnosti člen s γ, a tudíž na tento parametr není kladeno

žádné další omezení kromě $\gamma \in \mathbb{R}_+$. Vzhledem k předposledním dvěma nerovnicím ani vektory $\boldsymbol{\nu}_0$ a $\boldsymbol{\nu}_1$ nejsou omezeny žádnou podmínkou, a tak bude rovnost $\boldsymbol{\nu}_1 - \boldsymbol{\nu}_0 + \tau\boldsymbol{x} = \mathbf{0}$ platit pro libovolné $\boldsymbol{x} \in \mathbb{R}^n$ a $\tau \in \mathbb{R}_+$. Poslední nerovnost bude splněna, pokud bude výraz $\boldsymbol{x}^\top \boldsymbol{r} - r_e$ kladný. Vzhledem k tomu, že $\tau > 0$, jinak by dané omezení muselo být nutně porušeno, lze celou nerovnost vydělit τ a učinit limitní přechod pro $\tau \to +\infty$: $\boldsymbol{x}^\top \boldsymbol{r} - r_e \geq \frac{1}{\tau} \to 0$, což vede na úlohu (4.1).

Bohatství je pro jednoduchost voleno rovné jedné ($w = 1$). Jedná se pouze o škálovací faktor, kterým můžeme v případě potřeby optimální řešení $\boldsymbol{x}^*$, získané pro volbu $w = 1$, přenásobit konstantou a získat řešení pro obecnou volbu bohatství.

Ve všech třech optimalizačních úlohách vystupují parametry pravděpodobnostního rozdělení náhodného vektoru $\boldsymbol{\rho}$ ($\boldsymbol{\omega}$). V souladu s postupy uvedenými v Kapitole 2.1 lze dané parametry obsažené v účelové funkci i v pravděpodobnostním omezení nahradit jejich odhady $\hat{\boldsymbol{r}}$ a $\hat{\mathbb{V}}$. Tyto hodnoty budou určeny z historických dat metodou *exponenciálně vážených klouzavých průměrů* (viz Engle (2009) a Tsay (2005)), která dává větší váhu hodnotám z blízké minulosti před těmi časově vzdálenějšími. Odhad střední hodnoty $\boldsymbol{r}$ v čase n má pak pro vyrovnávací konstantu $\beta \in (0,1)$ tvar:

$$\hat{\boldsymbol{r}}_\beta(n) = (1-\beta)\sum_{i=0}^{n-1} \beta^i \boldsymbol{R}_{n-i} = (1-\beta)\boldsymbol{R}_n + \beta\hat{\boldsymbol{r}}_\beta(n-1), \tag{4.5}$$

kde $\boldsymbol{R}_n$ je vektor výnosů v čase n. Obdobně je také určen odhad kovarianční matice $\mathbb{V}$:

$$\begin{aligned} \hat{\mathbb{V}}_\beta(n) &= (1-\beta)\sum_{i=0}^{n-1} \beta^i (\boldsymbol{R}_{n-i} - \overline{\boldsymbol{R}}_n)(\boldsymbol{R}_{n-i} - \overline{\boldsymbol{R}}_n)^\top \\ &= (1-\beta)(\boldsymbol{R}_n - \overline{\boldsymbol{R}}_n)(\boldsymbol{R}_n - \overline{\boldsymbol{R}}_n)^\top + \beta\hat{\mathbb{V}}_\beta(n-1), \end{aligned} \tag{4.6}$$

kde $\overline{\boldsymbol{R}}_n$ je průměr výnosů přes sledované období.

4.2 Zpracování obrazu

Jedním ze základních problémů týkajících se zpracovávání obrazu je odstraňování šumu. Zašuměný černobílý obrázek ve formě $m \times n$ rozměrné matice může být reprezentován vektorem intenzit jednotlivých pixelů $\boldsymbol{f} \in [0,1]^{mn}$, kde 0 značí černou a 1 bílou barvu. V takovém případě je úkolem odseparovat od sebe původní obrázek $\boldsymbol{x} \in [0,1]^{mn}$ a šum, což je realizace náhodného vektoru $\boldsymbol{\omega} \in [-1,1]^{mn}$. Při tomto procesu se vychází z pozorování, že v obrázcích se vyskytují velké plochy, kde se intenzita mezi sousedními pixely mění pozvolna. Tyto plochy jsou odděleny ostrými hranicemi, kde dochází ke skokovité změně v intenzitě. Celkem to znamená, že obrázky obecně vykazují nízkou celkovou varianci:

$$TV(\boldsymbol{x}) = \sum_{1 \leq i < m} \sum_{1 \leq j < n} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2}, \tag{4.7}$$

kde $x_{i,j}$ označuje pixel nacházející se v i-tém řádku a j-tém sloupci obrázku $\boldsymbol{x}$. Na základě tohoto empirického pozorování byl v článku Goldfarb a Yin (2005) formulován následující optimalizační problém:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\omega}} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn},\, \boldsymbol{\omega} \in \mathbb{R}^{mn} \\ & \boldsymbol{x} + \boldsymbol{\omega} = \boldsymbol{f} \\ & \|\boldsymbol{\omega}\|_2 \leq K, \end{aligned} \tag{4.8}$$

kde $K > 0$ je zvolená konstanta. Tato úloha je konvexní, neboť normy jsou konvexními funkcemi a afinní funkce v omezení ve tvaru rovnosti tuto skutečnost neporuší. Nejčastějším předpokladem, který bývá kladen na šumovou složku v těchto modelech, je, že se skládá z nezávislých normálně rozdělených náhodných veličin o nulové střední hodnotě a známém fixním rozptylu σ^2.

Tyto předpoklady jsou však v mnohých případech zcela jistě porušeny. Například variabilita závisí na intenzitě v jednotlivých místech obrazu a bývá obecně vyšší v oblastech s vyšší intenzitou. Také nosič uvažovaného náhodného vektoru musí být omezený, neboť intenzity v zašuměném obrázku musí stále ležet v intervalu [0,1]. Pro tento případ, kdy neznáme přesné rozdělení náhodné veličiny $\boldsymbol{\omega}$, lze využít metody odvozené v Kapitole 3. Pro *pesimistickou variantu* by úloha měla tvar:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\omega}} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn} \\ & \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{x} + \boldsymbol{\omega} = \boldsymbol{f}] \geq 1 - \epsilon, \end{aligned} \tag{4.9}$$

kde rodina pravděpodobnostních rozdělení $\mathcal{P}$ je určena předpisem:

$$\mathcal{P} = \{\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{mn}) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\omega}] = \boldsymbol{\mu},\, \mathbb{E}_{\mathbb{P}}[\|\boldsymbol{\omega}\|_2] \leq \sigma\}.$$

Skutečnost, že velikost šumu v daném pixelu závisí na intenzitě, lze do modelu zanést, pokud bude střední hodnota náhodného vektoru $\boldsymbol{\omega}$ volena jako vhodná funkce zašuměného obrázku $\boldsymbol{\mu}(\boldsymbol{f})$. Takto definovaná rodina pravděpodobnostních rozdělení obsahuje i spojitá rozdělení, pro která platí $\mathbb{P}[\boldsymbol{x} + \boldsymbol{\omega} = \boldsymbol{f}] = 0$, a tudíž příslušné omezení:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{x} + \boldsymbol{\omega} = \boldsymbol{f}] \geq 1 - \epsilon$$

nemůže být pro $\epsilon \in (0,1)$ nikdy splněno.

V této aplikaci se proto zaměříme pouze na *optimistickou variantu*:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\omega}} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn} \\ & \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{x} + \boldsymbol{\omega} = \boldsymbol{f}] \geq 1 - \epsilon. \end{aligned} \tag{4.10}$$

Pro využití výsledků odvozených v Kapitole 3 je nutné výraz uvnitř pravděpodobnostního omezení vyjádřit ve tvaru nerovností:

$$\begin{aligned} \boldsymbol{\omega} &\leq \boldsymbol{f} - \boldsymbol{x} \\ -\boldsymbol{\omega} &\leq -(\boldsymbol{f} - \boldsymbol{x}). \end{aligned}$$

Tyto dvě soustavy nerovnic lze spojit do jedné velké soustavy:

$$A\boldsymbol{\omega} \leq A(\boldsymbol{f} - \boldsymbol{x}),$$

kde A je bloková matice o rozměru $2mn \times mn$ tvořena jednotkovými maticemi $\mathbb{I}_{mn\times mn}$:

$$A = \begin{pmatrix} \mathbb{I} \\ -\mathbb{I} \end{pmatrix}.$$

Tím jsou omezení převedeny do tvaru neostrých nerovností. Převod na ostré lze učinit snadno vložením konstantního vektoru $\delta\mathbb{1}$, kde $\mathbb{1}$ je vektor jedniček a $\delta > 0$ je malé:

$$A\boldsymbol{\omega} < A(\boldsymbol{f} - \boldsymbol{x}) + \delta\mathbb{1}.$$

Nyní je úloha v požadovaném tvaru:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\omega}} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn} \\ & \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\omega} < \boldsymbol{u}(\boldsymbol{x})] \geq 1 - \epsilon, \end{aligned} \tag{4.11}$$

kde $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T} = A$ a $\boldsymbol{u}(\boldsymbol{x}) = A(\boldsymbol{f} - \boldsymbol{x}) + \delta\mathbb{1}$. Jelikož je vybraná disperzní funkce (l^2-norma) pozitivně homogenní, množina $\boldsymbol{\Omega} = \mathbb{R}^{mn}$ je konvexní uzavřený kužel, matice $\boldsymbol{T} = A$ je konstantní a kužel $\mathcal{D} = \mathbb{R}_+$, lze pravděpodobnostní omezení úlohy (4.11) dle výsledků z Kapitoly 3.3 přepsat jako:

$$\begin{aligned} & \boldsymbol{\nu}_0,\, \boldsymbol{\nu}_1 \in \mathbb{R}^{mn} \\ & A\boldsymbol{\nu}_0 \leq (1-\epsilon)(A(\boldsymbol{f} - \boldsymbol{x}) + \delta\mathbb{1}) \\ & \sum_{i\in\{0,1\}} \boldsymbol{\nu}_i = \boldsymbol{\mu}(\boldsymbol{f}) \\ & \sum_{i\in\{0,1\}} \|\boldsymbol{\nu}_i\|_2 \leq \sigma \end{aligned}$$

a celou úlohu lze napsat ve tvaru:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\omega}} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn},\, \boldsymbol{\nu}_0,\, \boldsymbol{\nu}_1 \in \mathbb{R}^{mn} \\ & \boldsymbol{\nu}_0,\, \boldsymbol{\nu}_1 \in \mathbb{R}^{mn} \\ & A\boldsymbol{\nu}_0 \leq (1-\epsilon)(A(\boldsymbol{f} - \boldsymbol{x}) + \delta\mathbb{1}) \\ & \sum_{i\in\{0,1\}} \boldsymbol{\nu}_i = \boldsymbol{\mu}(\boldsymbol{f}) \\ & \sum_{i\in\{0,1\}} \|\boldsymbol{\nu}_i\|_2 \leq \sigma. \end{aligned} \tag{4.12}$$

Dále je potřeba z úlohy eliminovat konstantu δ, která do ní bylo uměle přidána při jejím převodu do požadovaného tvaru (4.11). Tento proces je obsahem následujícího lemmatu.

Lemma 4.1. *Označme optimální řešení úlohy (4.12) jako* $(\boldsymbol{x}_\delta, \boldsymbol{\nu}_{0,\delta}, \boldsymbol{\nu}_{1,\delta})$. *Potom pro* $\delta \to 0$ *je tato posloupnost omezená a každá konvergentní podposloupnost je řešením:*

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\nu}_0,\boldsymbol{\nu}_1} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn},\ \boldsymbol{\nu}_0,\ \boldsymbol{\nu}_1 \in \mathbb{R}^{mn} \\ & \boldsymbol{\nu}_0 = (1-\epsilon)(\boldsymbol{f} - \boldsymbol{x}) \\ & \sum_{i\in\{0,1\}} \boldsymbol{\nu}_i = \boldsymbol{\mu}(\boldsymbol{f}) \\ & \sum_{i\in\{0,1\}} \|\boldsymbol{\nu}_i\|_2 \leq \sigma. \end{aligned} \tag{4.13}$$

Důkaz. Nejprve ukažme, že optimální řešení úlohy (4.12) existuje. To plyne ze skutečnosti, že účelová funkce $TV(\boldsymbol{x})$ je spojitá a množina přípustných řešení je uzavřená a omezená (kompaktní). Dále vyberme konvergentní podposloupnost ve tvaru $(\boldsymbol{x}_{\delta_k}, \boldsymbol{\nu}_{0,\delta_k}, \boldsymbol{\nu}_{1,\delta_k})$ a ukážeme, že její limita $(\boldsymbol{x}^*,\boldsymbol{\nu}_0^*,\boldsymbol{\nu}_1^*)$ pro $k \to +\infty$ ($\delta_k \to 0$) je přípustným řešením úlohy (4.13). Množina přípustných řešení úlohy (4.12) je kompaktní, a tudíž limita konvergentní posloupnosti bude prvkem stejné množiny. Z toho plyne splnění omezení $\boldsymbol{x}^* \in [0,1]^{mn}$, $\sum_{i\in\{0,1\}} \boldsymbol{\nu}_i^* = \boldsymbol{\mu}(\boldsymbol{f})$ a $\sum_{i\in\{0,1\}} \|\boldsymbol{\nu}_i^*\|_2 \leq \sigma$. Splnění zbývající podmínky vyplývá z toho, že každý bod vybrané podposlouplosti $(\boldsymbol{x}_{\delta_k}, \boldsymbol{\nu}_{0,\delta_k}, \boldsymbol{\nu}_{1,\delta_k})$ leží v množině určené nerovnostmi:

$$\begin{aligned} \boldsymbol{\nu}_{\delta_k} &\leq (1-\epsilon)(\boldsymbol{f} - \boldsymbol{x}_{\delta_k} + \delta_k \mathbb{1}) \\ -\boldsymbol{\nu}_{\delta_k} &\leq -(1-\epsilon)(\boldsymbol{f} - \boldsymbol{x}_{\delta_k} - \delta_k \mathbb{1}). \end{aligned}$$

Pokud má limita podposloupnosti zůstat v množině přípustných řešení, musí ležet v průniku těchto množin, což je množina:

$$\boldsymbol{\nu} = (1-\epsilon)(\boldsymbol{f} - \boldsymbol{x}),$$

a tedy bod $(\boldsymbol{x}^*,\boldsymbol{\nu}_0^*,\boldsymbol{\nu}_1^*)$ je přípustným řešením úlohy (4.13).

Nyní je potřeba ukázat, že $\lim_{k\to+\infty} TV(\boldsymbol{x}_{\delta_k}) = TV(\bar{\boldsymbol{x}})$, kde $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\nu}}_0, \bar{\boldsymbol{\nu}}_1)$ je optimální řešení problému (4.13). Začneme s důkazem nerovnosti $\lim_{k\to+\infty} TV(\boldsymbol{x}_{\delta_k}) \leq TV(\bar{\boldsymbol{x}})$, která platí díky tomu, že množina přípustných řešení úlohy (4.12) je pro každé δ_k větší než množina přípustných řešení úlohy (4.13). Druhá nerovnost plyne z toho, že bod $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\nu}}_0, \bar{\boldsymbol{\nu}}_1)$ je optimálním řešením problému (4.13), a tudíž:

$$TV(\bar{\boldsymbol{x}}) \leq TV(\boldsymbol{x}^*) = \lim_{k\to+\infty} TV(\boldsymbol{x}_{\delta_k}).$$

Tím je ověřeno, že limita optimálních řešení úlohy (4.12) $(\boldsymbol{x}_\delta, \boldsymbol{\nu}_{0,\delta}, \boldsymbol{\nu}_{1,\delta})$ pro $\delta \to 0$ je optimálním řešením úlohy (4.13). □

Problém (4.13) zůstává konvexní, neboť l^2-norma, a tedy i funkce $TV(\boldsymbol{x})$, je konvexní a ostatní podmínky jsou ve tvaru lineárních omezení. Vztah mezi touto úlohou a původním problémem (4.8) lze odhalit při volbě $\boldsymbol{\mu}(\boldsymbol{f}) = \boldsymbol{0}$. V takovém případě z omezení $\sum_{i\in\{0,1\}} \boldsymbol{\nu}_i = \boldsymbol{\mu}(\boldsymbol{f})$ vyplývá, že $\boldsymbol{\nu}_0 = -\boldsymbol{\nu}_1$. Za použití substituce $\boldsymbol{\omega} = \frac{\boldsymbol{\nu}_0}{1-\epsilon}$ lze problém (4.13) vyjádřit ve tvaru:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\omega}} \quad & TV(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \in [0,1]^{mn},\ \boldsymbol{\omega} \in \mathbb{R}^{mn} \\ & \boldsymbol{x} + \boldsymbol{\omega} = \boldsymbol{f} \\ & \|\boldsymbol{\omega}\|_2 \leq \frac{\sigma}{2(1-\epsilon)}. \end{aligned}$$

Úloha (4.13) tedy přechází na původní problém (4.8) s parametrem:

$$K = \frac{\sigma}{2(1-\epsilon)}. \tag{4.14}$$

Pro eliminaci základních nedostatků standardní metody reprezentované úlohou (4.8) je nezbytné volit střední hodnotu šumové složky nenulovou a také postihnout její závislost na intenzitě v jednotlivých pixelech obrázku $\boldsymbol{f}$. Nejjednodušší lineární závislost může vypadat takto:

$$\boldsymbol{\mu}(\boldsymbol{f}) = b(\mathbb{1} - 2\boldsymbol{f}). \tag{4.15}$$

Tento model odráží fakt, že v příliš temných, nebo světlých oblastech obrázku nemůže mít šumová složka nulovou střední hodnotu, neboť intenzita v každém pixelu originálního i zašuměného obrázku musí být v intervalu [0,1]. Význam parametru $b > 0$ je takový, že udává rozmezí pro střední hodnoty šumové složky obrázku. V nejtemnějších oblastech je střední hodnota rovna b a v nejsvětlejších místech má naopak střední hodnota velikost $-b$. Výběru vhodných hodnot parametrů b a σ bude věnována výpočetní část.

5. Výpočetní část

Pro řešení vybraných optimalizačních úloh byl zvolen software *Mathlab* ve verzi *R2017b* spolu s programy *Yalmip* a *CVX*, které slouží k přípravě vstupů do formy vyžadované použitými řešiteli optimalizačních úloh. Zdrojové kódy použité při řešení jednotlivých úloh jsou k dispozici jako elektronické přílohy.

5.1 Optimalizace portfolia

Vybrané modely z předchozí části budou nyní aplikovány na portfolio tvořeno 10 vybranými akciovými tituly. Pro výpočet odhadů střední hodnoty výnosů akcií $\boldsymbol{r}$ a kovarianční matice $\mathbb{V}$ byla použita data získaná ze serveru *Yahoo! Finance* za období od 7.4.2017 do 7.4.2018 (celkem 250 pozorování). Odhady byly počítány metodou *exponenciálně vážených klouzavých průměrů* s volbou vyrovnávací konstanty $\beta = 0{,}9$ dle doporučení v knize Cipra (2008). Denní výnos akcie v čase t R_t byl určen z denních uzavíracích kurzů akcií očištěných o dividendu p_t jako:

$$R_i = \frac{p_t - p_{t-1}}{p_{t-1}} 100[\%].$$

Vybrané akcie spolu s očekávaným denním výnosem a rizikem (směrodatnou odchylkou) jsou uvedeny v tabulce 5.1.

Tabulka 5.1: Vybrané akciové tituly

Akcie	Eminent	Očekávaný výnos [%]	Riziko [%]
OXY	Occidental Petroleum Corp.	0,30	1,87
LAC	Lithium Americas Corp.	-0,46	4,46
PFE	Pfizer Inc.	-0,15	1,35
DIS	The Walt Disney Comp.	-0,08	1,42
C	Citigroup Inc.	-0,28	1,74
UL	Unilever PLC	0,42	1,39
DAL	Delta Air Lines, Inc.	-0,23	2,09
ADS.DE	Adidas AG	0,63	2,20
SIE.DE	Siemens AG	0,05	1,56
VOW.DE	Volkswagen AG	0,24	1,73

Na portfolio těchto akcií byly aplikovány všechny tři optimalizační postupy. Níže uvedené výpočty byly učiněny pro velikost bohatství $w = 1$, $r_e = 0{,}5$ a $\epsilon = 0{,}1$. Parametr σ byl vybírán v rozmezí od 0 do $\sqrt{\mathrm{Tr}\,\mathbb{V}} = 6{,}8$. Výsledky analýz v závislosti na volbě tohoto parametru jsou znázorněny na obrázcích 5.1, 5.2 a 5.3. Do grafů je přidána dolní mez úlohy (4.1), která vznikne jako řešení této úlohy bez podmínky na minimální očekávaný výnos ($\boldsymbol{r}^\top \boldsymbol{x} \geq r_e$). V grafech si můžeme všimnout, že pro volbu $\sigma = 0$ dávají všechny přístupy stejný výsledek, což potvrzuje úvahy uvedené v předchozí kapitole.

V případě *pesimistického scénáře* s rostoucím parametrem σ roste také výnos a riziko vybraného portfolia. To souvisí s faktem, že při *pesimistické variantě* jsme nuceni do portfolia zahrnout i akcie, které mají vyšší kladný, ale i záporný

výnos (*prodeje na krátko*) a které jsou současně více rizikové. Tento efekt lze pozorovat na váhách akciích LAC a ADS.DE v tabulce 5.2. Pro *pesimistický scénář* neexistuje žádná horní hranice, a tudíž aby bylo zajištěno splnění podmínky na minimální očekávaný výnos (viz úloha (4.2)), je třeba v portfoliu zvyšovat zastoupení nejvýnosnějších (a nejrizikovějších) akcií. Riziko takto vytvořeného portfolia pak se zvyšováním velikosti parametru σ enormně narůstá. Na obrázcích 5.1 až 5.3 není zachycen celý průběh závislostí v daném rozsahu hodnot σ. To je vedle zachování přehlednosti obrázků také dáno tím, že zhruba od hodnoty $\sigma = 0{,}2$ již nebylo možné nalézt přípustné řešení. Vzhledem k praktickému použití získaných výsledků však ani při nižších hodnotách parametru σ, které nejsou znázorněny na obrázcích, nepřicházíme o žádná zajímavá řešení, neboť se jedná o portfolia, jejichž rizika jsou výrazně vyšší než riziko portfolia získaného *Markowitzovým přístupem*.

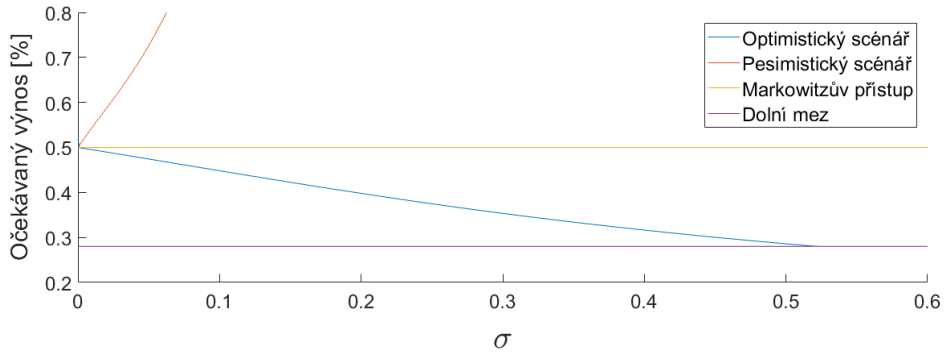
Naopak pro *optimistický scénář* platí, že do portfolia lze zahrnout i akcie s nižším výnosem (a rizikem), protože věříme, že jejich výnos bude vyšší než očekávaný. V tomto případě je také očekávaný výnos nižší než předepsaná hodnota 0,5 % (viz obrázek 5.1). Pokud je hodnota parametru σ zvyšována, dochází k rozšiřování rodiny pravděpodobnostních rozdělení, ve které se nachází pravděpodobnostní rozdělení náhodného vektoru $\boldsymbol{\omega}$ popisujícího perturbace od střední hodnoty vektoru výnosů akcií v portfoliu. Při jisté hodnotě σ, v tomto konkrétním případě zhruba 0,53, je rodina pravděpodobnostních rozdělení tak bohatá, že je možné současně splnit *optimistickou variantu* pravděpodobnostního omezení (3.19) a také dosáhnout zmíněné dolní hranice účelové funkce. Poté již zvyšování hodnoty parametru σ nijak neovlivní optimální řešení dané úlohy, které zůstává stejné.

Na obrázeku 5.3 jsou vyneseny hodnoty skutečných výnosů portfolií vzniklých třemi různými přístupy pro následující obchodní den (9.4.2018). Zde je vidět, že pro daný soubor akcií dává nejlepší výsledek metoda *optimistických scénářů*, kdy byly pro parametry σ v intervalu od 0 do zhruba 2,8 tímto postupem sestaveny portfolia s nejvyšším reálným výnosem a nižším rizikem v porovnání s portfoliem vzešlým ze standardního *Markowitzova přístupu*.

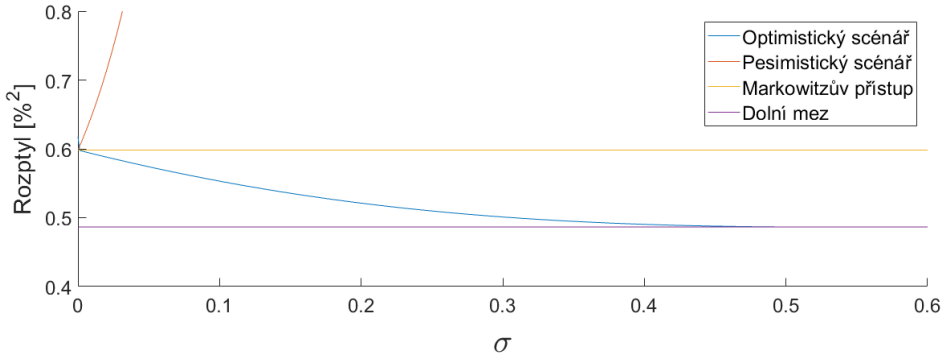
Poslední porovnání použitých přístupů je znázorněno na obrázku 5.4, kde jsou vyneseny závislosti rozptylů portfolií získaných metodami *optimistických* a *pesimistických scénářů* na jejich očekávaných výnosech. Pro *Markowitzův přístup* je zde závislost rozptylů portfólií na předepsaném výnosu r_e, který je však od jisté hranice rovný výnosu očekávanému. Průběh obou závislostí je velice podobný. *Optimistický scénář* nabývá nižších hodnot očekávaných výnosů i rozptylu a v bodě zvoleného předepsaného výnosu 0,5 % se napojuje na *pesimistický scénář*, pro nějž rozptyl portfolia stoupá s narůstajícím očekávaným výnosem strměji než pro *Markowitzův přístup*.

Tabulka 5.2: Porovnání vah akcií v portfoliu pro různé přístupy

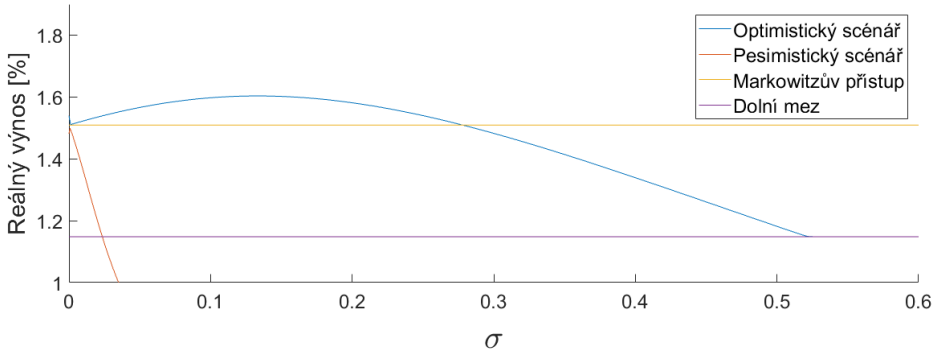
	Markowitzův přístup	Pesimistický scénář $\sigma = 0{,}05$	$\sigma = 0{,}1$	Optimistický scénář $\sigma = 0{,}3$	$\sigma = 0{,}6$
OXY	0,47	0,50	0,63	0,39	0,31
LAC	-0,07	-0,09	-0,18	-0,05	-0,03
PFE	0,36	0,05	-0,16	0,49	0,48
DIS	-0,22	-0,06	-0,04	-0,19	-0,10
C	-0,21	-0,27	-0,42	-0,19	-0,18
UL	0,47	0,62	0,78	0,34	0,28
DAL	-0,17	-0,15	-0,18	-0,15	-0,13
ADS.DE	0,12	0,24	0,50	0,07	0,04
SIE.DE	0,34	0,11	-0,01	0,39	0,33
VOW.DE	-0,10	0,05	0,09	-0,09	-0,01



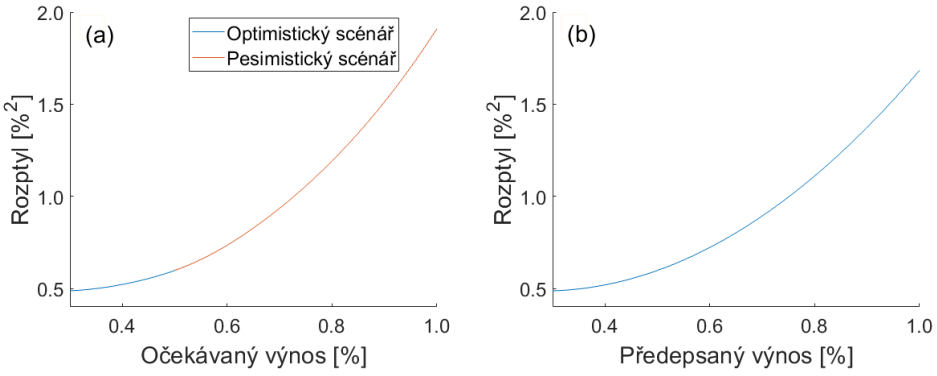
Obrázek 5.1: Závislost očekávaného výnosu portfolia na volbě parametru σ



Obrázek 5.2: Závislost rozptylu portfolia na volbě parametru σ



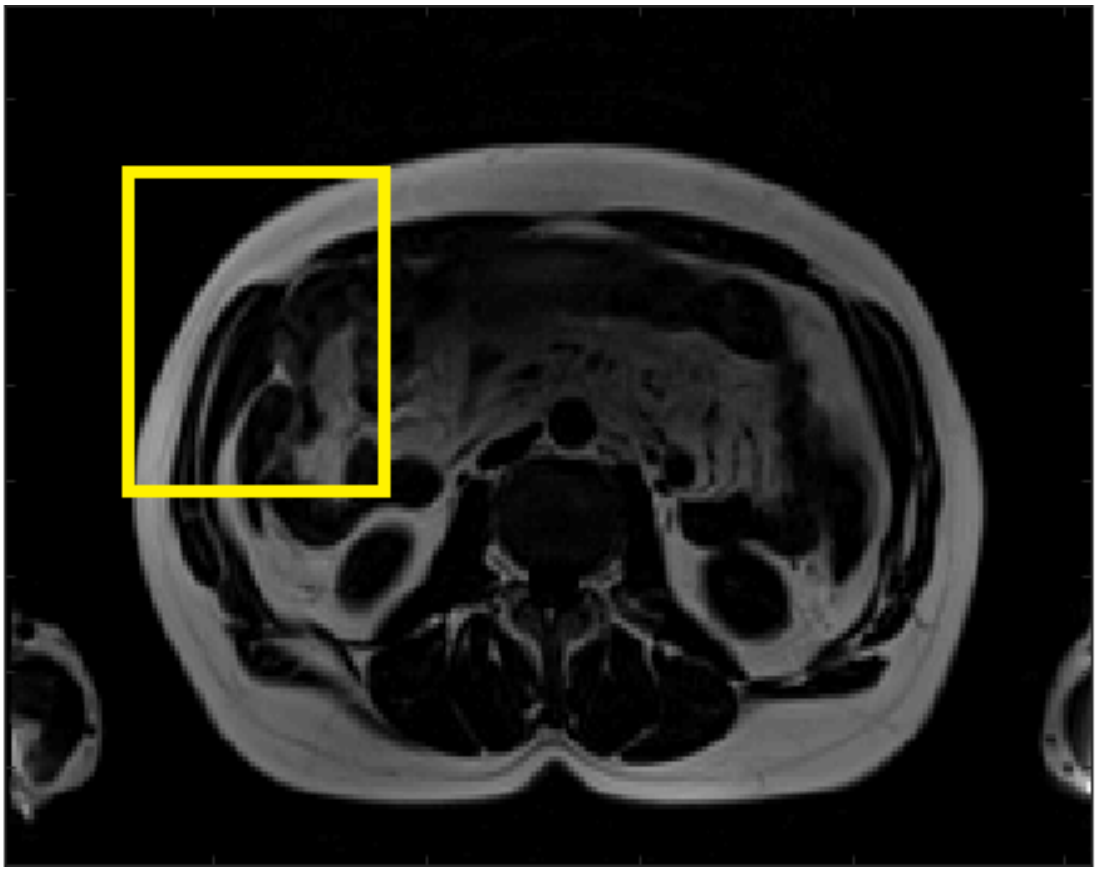
Obrázek 5.3: Závislost reálného výnosu portfolia na volbě parametru σ



Obrázek 5.4: Závislost rozptylu portfolia na očekávaném výnosu pro *pesimistický* a *optimistický scénář* (a) a na předepsaném výnosu pro *Markowitzův přístup* (b)

5.2 Zpracování obrazu

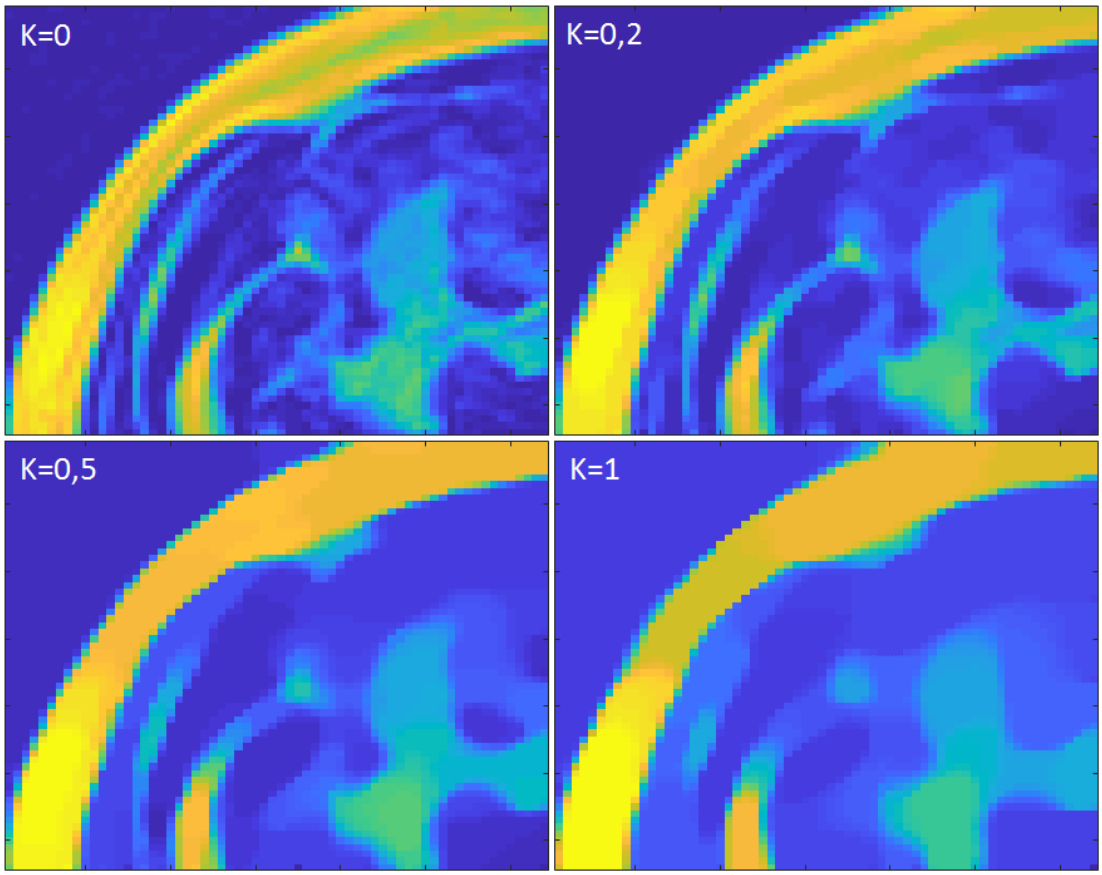
V této části budou postupy odvozené v Kapitole 4.2 aplikovány na reálná data. Konkrétně na obrázek z magnetické rezonance zobrazující transverzální řez lidského těla (viz obrázek 5.5). Hlavním cílem je porovnat standardní přístup reprezentovaný úlohou (4.8) s *optimistický scénář* (4.13). Volba parametru ϵ je fixní na hodnotě 0,1 a ostatní parametry budou nastavovány tak, aby bylo dosaženo co nejlepší kvality obrázku. V tomto případě neexistuje žádné objektivní měřítko, a tak rozhodování a výběr nejlepších parametrů budou založeny na subjektivních kritériích. Vzhledem k výpočetním možnostem použitého softwaru byla analyzována pouze žlutě označená část zmíněného obrázku 5.5. V rámci standardního přístupu byl nejvhodnější parametr K z úlohy (4.8) vybírán v rozmezí od 0 do 1. Na obrázku 5.6 jsou znázorněny výsledky pro tři konkrétní volby parametru K. Obrázky jsou vykresleny v barevné škále, aby byly lépe znatelné drobné rozdíly. Původní obrázek odpovídá volbě $K = 0$, jež obsahuje mnoho bodů, které svou intenzitou nezapadají do okolních struktur. Tento efekt může být zmírněn metodami sloužícími k odstranění šumové složky. Na dalších třech obrázcích je



Obrázek 5.5: Obrázek z magnetické rezonance se žlutě ohraničeným výsekem, na kterém budou prováděny další analýzy

demonstrován efekt parametru K. Při jeho zvyšování se povoluje větší rozdíl mezi původním obrázkem $\boldsymbol{f}$ a vyhlazeným obrázkem $\boldsymbol{x}$ (viz problém (4.8)). Tím se rozšiřuje množina přípustných řešení, a tedy optimální hodnota účelové funkce $TV(\boldsymbol{x})$ klesá. To se vizuálně projeví tím, že je daný obrázek méně ostrý. Při tomto procesu dochází jednak k odstraňování šumu, ale také ke ztrátě kontrastu. Je tedy

potřeba najít kompromis mezi těmito dvěma protichůdným efekty. Proto byla zvolena hodnota parametru $K = 0{,}2$, při níž ještě nedošlo k zásadnímu rozmazání přítomných struktur, což už pro vyšší volby parametru K neplatí.

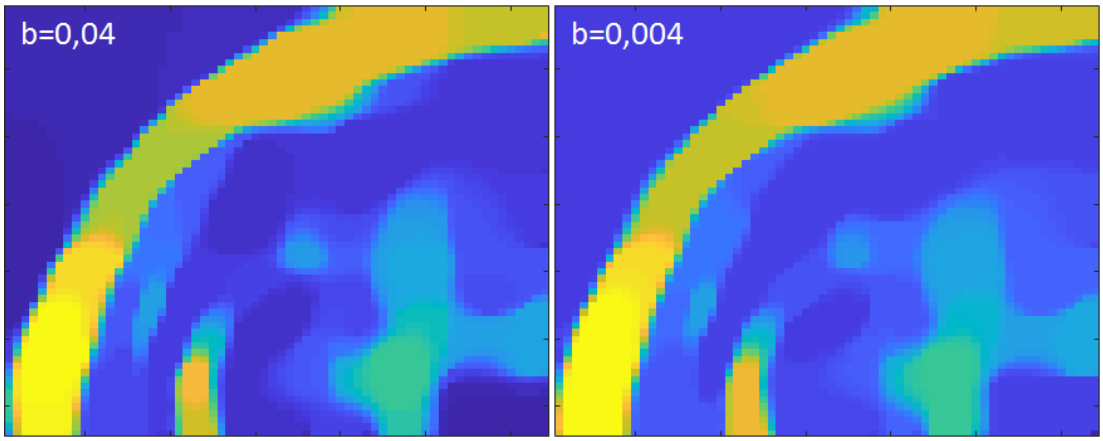


Obrázek 5.6: Standardní přístup

Metoda *optimistických scénářů* zahrnuje, při pevném ϵ, dva volné parametry. Jejich hodnoty byly vybírány v rozmezí od 0 po 2 v případě parametru σ. Horní mez pro parametr b závisí na volbě σ. Pro $\sigma = 2$ bylo možné volit parametr b až do hodnoty 0,04. V případě vyšší hodnoty parametru b již neexistuje přípustné řešení. Vliv tohoto parametru je znázorněn na obrázku 5.7. Pro tento účel byla záměrně volena vyšší hodnota parametru σ, díky čemuž je množina možných parametrů b větší, a tudíž je jeho efekt znatelnější. Ten spočívá ve zmírnění ztráty kontrastu na hranách, který je například vidět na struktuře v pravé dolní části obrázku. Pokud při fixní volbě parametru σ dojde ke zvýšení hodnoty parametru b, tak se začnou jednolité celistvé plochy v obrázku rozpadat na menší s drobnými rozdíly v intenzitách.

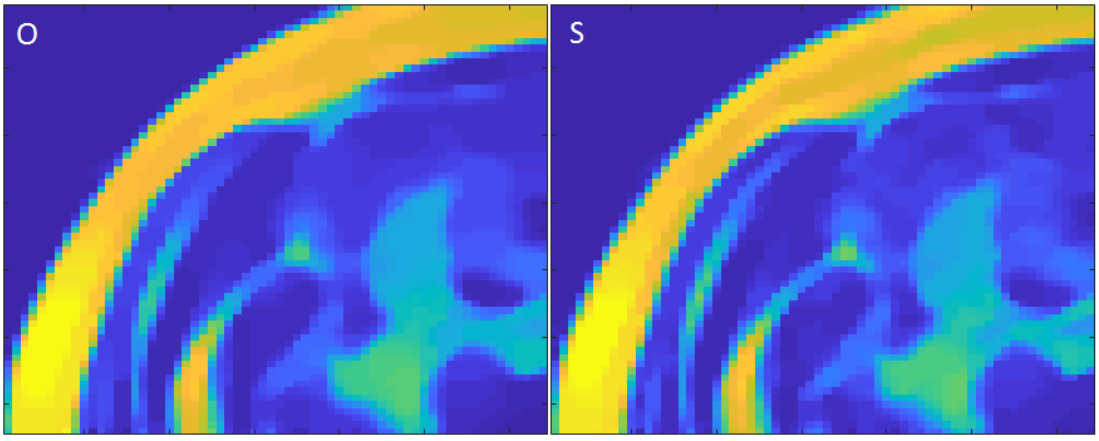
Tabulka 5.3: Výběr optimálních parametrů

	Standardní přístup	Optimistický scénář	
Parametry	K	σ	b
Hodnoty	0,2	0,6	0,01



Obrázek 5.7: Vliv parametru b na vyhlazování obrázku pro $\sigma = 2$

Porovnání obou metod pro nejlepší vybrané parametry, jejichž hodnoty jsou uvedeny v tabulce 5.3, je na obrázku 5.8. Zajímavé je, že pro *optimistický scénář* byla zvolena hodnota $\sigma = 0{,}6$, která je vyšší než hodnota spočítaná podle vzorce (4.14), jež udává vztah mezi parametrem K pro standardní přístup a parametrem σ z *optimistického scénáře* za předpokladu nulové střední hodnoty šumové složky ($b = 0$). To znamená, že pro model s nenulovým parametrem b lze použít větší vyhlazení (vyšší σ), čímž v obrázku vzniknou větší plochy konstantních intenzit, při malé ztrátě ostrosti (viz obrázek 5.8). To je dáno výběrem funkce $\boldsymbol{\mu}(\boldsymbol{f})$ (viz vzorec (4.15)), která v případě pixelu s vysokou intenzitou v obrázku $\boldsymbol{f}$ preferuje vysokou hodnotu intenzity i ve vyhlazeném obrázku $\boldsymbol{x}$. Protože platí, že $\boldsymbol{x} = \boldsymbol{f} - \boldsymbol{\omega}$ a střední hodnota $\boldsymbol{\mu}(\boldsymbol{f})$ je po složkách klesající funkce intenzit v obrázku $\boldsymbol{f}$. Na rozhraní dvou míst s (výrazně) rozdílnou intenzitou se tyto vlivy sčítají a brání tím rozmazání dané hranice. Tento efekt je však malý a vždy pouze koriguje rozmazání dané velikostí parametru σ. Pomocí vhodné volby parametrů je možné docílit většího, či menšího počtu celistvých ploch v obrázku za stejného kontrastu na hranách.



Obrázek 5.8: Porovnání *optimistického scénáře* (O) se standardním přístupem (S)

Závěr

Obsahem diplomové práce bylo nalezení vhodných přístupů pro řešení optimalizačních úloh s pravděpodobnostními omezeními bez znalosti pravděpodobnostního rozdělení. Tento problém byl formulován v podobě dvou scénářů. *Pesimistický scénář* popisuje situaci, kdy k danému pravděpodobnostnímu omezení přistupujeme tak, že z množiny možných pravděpodobnostních rozdělení vybíráme to nejméně příznivé. Při *optimistickém scénáři* naopak hledáme pravděpodobnostní rozdělení, které nám poskytne co možná největší množinu přípustných řešení. V teoretické části byly tyto úlohy s pravděpodobnostními omezeními transformovány na jednodušší deterministické úlohy.

Následně byly odvozené přístupy aplikovány na praktické problémy. Nejprve jsme se zaměřili na problém *optimalizace portfolia*, kdy se investor snaží alokovat své bohatství mezi portfolia aktiv tak, aby minimalizoval riziko za podmínky dosažení alespoň předepsaného výnosu. Zde byly metody scénářů porovnány na reálných datech s *Markowitzovým přístupem*, kde je s využitím odhadů parametrů pravděpodobnostního rozdělení výnosů akcií řešena deterministická optimalizační úloha. Při použití *pesimistického scénáře* je investor nucen více riskovat, aby zajistil splnění požadavku na předepsaný výnos. Tudíž musí investovat větší část svého bohatství do akcií s vyšším výnosem a obvykle také s větším rizikem. Čím je pesimismus investora, vyjádřený pomocí parametrů modelu, větší, tím roste očekávaný výnos a také enormně narůstá riziko takto sestaveného portfolia.

Přijatelnějším přístupem se zdá být *optimistický scénář*. Očekávaná rizika portfolií nalezených tímto postupem jsou menší nebo rovna riziku portfolia sestaveného pomocí *Markowitzova přístupu*. V případě rozšiřování rodiny pravděpodobnostních rozdělení, ze které vybíráme, dochází ke snižování očekávaného výnosu (pod předepsanou úroveň) spolu s rizikem. Tento pokles se zastaví v momentě, kdy je rodina pravděpodobnostních rozdělení dostatečně velká na to, aby bylo možné dosáhnout globálního minima pro riziko portfolia s daným souborem akcií a bohatstvím investora. Toto portfolio odpovídá optimálnímu řešení úlohy na minimalizaci rizika bez podmínky na minimální předepsaný výnos. Nejlepším přístupem se zdá být metoda *optimictických scénářů*. V případě uvážené volby parametru vystihujícího míru optimismu investora bylo dosaženo vyššího reálného výnosu a současně nižšího očekávaného rizika než v případě *Markowitzova přístupu*.

Druhým problémem, na který byly získané přístupy použity, je *zpracování obrazu*. Konkrétně odstraňování šumu z obrázku naměřeného pomocí magnetické rezonance. *Pesimistický přístup* byl hned na začátku zavržen, neboť v případě volby nejhoršího možného pravděpodobnostního rozdělení z dané rodiny, bychom nikdy nemohli zaručit splnění pravděpodobnostního omezení. Během implementace *optimistického scénáře* na tento konkrétní problém se ukázalo, že se jedná o rozšíření standardní metody, kdy je zašuměný obrázek rozkládán na původní obrázek bez šumu a samotnou šumovou složku. Tato metoda v sobě obsahuje předpoklad o nulové střední hodnotě šumové složky. Tento předpoklad není zcela reálný, protože například v místech obrázku s maximální možnou intenzitou může šumová složka nabývat pouze záporných hodnot, protože ani v zašuměném obrázku nelze překročit hranici maximální intenzity. Tento efekt lze do úlohy zahr-

nout pomocí *optimistických scénářů*. Střední hodnota šumové složky pak může záviset na intenzitě v daném pixelu. Započítáním tohoto efektu do optimalizační úlohy je možné mírně zmírnit ztrátu kontrastu na rozhraní dvou oblastí s výrazně odlišnými intenzitami, což je nechtěný efekt všech metod sloužících k odstraňování šumu z obrázků.

Literatura

Beck, A. a Ben-Tal, A. (2009). Duality in robust optimization: primal worst equals dual best. *Operations Research Letters*, **37**(1), 1–6.

Boyd, S. a Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Cipra, T. (2008). *Finanční ekonometrie*. Ekopress.

Cordeau, J.-F., Laporte, G., Savelsbergh, M. W. a Vigo, D. (2007). Vehicle routing. *Handbooks in operations research and management science*, **14**, 367–428.

Dert, C. a Oldenkamp, B. (2000). Optimal guaranteed return portfolios and the casino effect. *Operations Research*, **48**(5), 768–775.

Engle, R. (2009). *Anticipating correlations: a new paradigm for risk management*. Princeton University Press.

Goldfarb, D. a Yin, W. (2005). Second-order cone programming methods for total variation-based image restoration. *SIAM Journal on Scientific Computing*, **27**(2), 622–645.

Hanasusanto, G. A., Roitch, V., Kuhn, D. a Wiesemann, W. (2017). Ambiguous joint chance constraints under mean and dispersion information. *Operations Research*, **65**(3), 751–767.

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, **7**(1), 77–91.

Prékopa, A. (1971). Logarithmic concave measures with application to stochastic programming. *Acta Scientiarum Mathematicarum*, **32**, 301–316.

Rockafellar, R. T. (1997). Convex analysis. princeton landmarks in mathematics.

Rockafellar, R. T. a Wets, R. J.-B. (2009). *Variational analysis*, volume 317. Springer Science & Business Media.

Shen, S., Smith, J. C. a Ahmed, S. (2010). Expectation and chance-constrained models and algorithms for insuring critical paths. *Management Science*, **56**(10), 1794–1814.

Tsay, R. S. (2005). *Analysis of financial time series*, volume 543. John Wiley & Sons.

Wang, J. (2007). The β-reliable median on a network with discrete probabilistic demand weights. *Operations research*, **55**(5), 966–975.