MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Marek Šešulka

Exaktní penalizace v optimalizaci

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Branda Martin, Ph.D
Studijní program: Matematika
Studijní obor: Obecná matematika

Praha 2019

Prohlašuji, že jsem tuto bakalářskou 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 .... dne .... Podpis autora

Děkuji svému vedoucímu bakalářské práce RNDr. Martinovi Brandovi, Ph.D za jeho odborné vedení a připomínky k práci. Dále bych chtěl poděkovat moji rodině, že ve mě věřili po celou dobu studia i v těch momentech kdy jsem byl sám na vážkách. V neposlední řadě děkuji mým dvěma nejbližším spolužákům Stanislavovi Šípkovi a Dominiku Rózsahegyimu, kteří mi vždy pomohli a podpořili.

Název práce: Exaktní penalizace v optimalizaci

Autor: Marek Šešulka

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Branda Martin, Ph.D, Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce se věnuje jednomu z možných přístupů řešení nelineárních optimalizační úloh a to převedením na úlohu hledání volného extrému, kde jsou omezení přenesena do účelové funkce pomocí takzvané penalizační funkce. Představíme Metodu vnějšího bodu a příslušný algoritmus pro řešení daného problému. Práce se dále zabývá exaktními penalizačními funkcemi, které nevyžadují limitní přiblížení penalizačního parametru k nekonečnu. Poté se zaobíráme celočíselným binárním nelineárním programováním, kde je uvedeno několik vhodných penalizačních funkcí pro řešení tohoto typu úloh. V numerické části se práce věnuje minimalizaci rizika při zadaném minimálním očekávaném výnosu portfolia s omezeným počtem aktiv. Sleduje vliv změny penalizačního parametru na výsledky deseti různých minimalizací rizika portfolií.

Klíčová slova: exaktní penalizace, nelineární programování, penalizační metoda, binární programování

Title: Exact penalization in optimization

Author: Marek Šešulka

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Branda Martin, Ph.D, Department of Probability and Mathematical Statistics

Abstract: This thesis deals with one of the possible different approaches to solving nonlinear optimization problems by convertion to finding non-bounded extrema of function, where constrains are transfered to objective function via penalty function. We will introduce exterior penalty function method and appropriate algorithm for solving this type for problems. The thesis also deals with exact penalty functions, which do not requires limit approximation of the penalty parameter to infinity. Then we deal with integer binary nonlinear progamming, where several suitable penalty functions are presented to solve this type of problem. In the numerical part, the thesis deals with the minimization of risk at the specifed minimum expected return on the sparse portfolio. We observe the effect of changing the penalty parameter on the results of ten different minimization problems calculating risk of sparsity portfolios.

Keywords: exact penalization, nonlinear programming, penalty function method, binary programming

Obsah

Úvod

Jedním z možných přístupů k řešení nelineárních úloh, je použití vhodné penalizační funkce, která je přidána do funkce účelové tak, aby jakékoliv porušení daných omezení bylo penalizováno velkou hodnotou. Poté tedy řešíme úlohu hledaní volného extrému, jejichž řešení spočívá v postupném zvyšování penalizačního parametru.

V teoretické částí práce si nejprve představíme nelineární optimalizační problém a jeho příslušný penalizační problém. Poté se seznámíme s Metodou vnějšího bodu a také s algoritmem pro řešení penalizačního problému. Algoritmus demonstrujeme na jednoduchém příkladu.

V další části se nejprve budeme zabývat exaktními penalizačními funkcemi, konkrétně absolutní penalizační funkcí. Exaktní penalizační funkce nevyžaduje nekonečné zvyšování penalizačního parametru, oproti Metodě vnějšího bodu, nýbrž funguje pro rozumný konečný penalizační parametr, který umíme určit. Opět budeme demonstrovat na příkladu. V podkapitole si zadefinujeme celočíselný binární nelineární optimalizační problém a jeho příslušný penalizační problém. Pak si ukážeme několik jednoduchých funkcí, jež daný binární penalizační problém řeší.

V numerické části se budeme věnovat problému minimalizace rizika při zadaném minimálním očekávaném výnosu portfolia s předem omezeným počtem aktiv. Jedná se o modifikovaný Markowitzův model. Daný problém převedeme na binární nelineární optimalizační problém a pomocí předchozích znalostí penalizačních funkcí, ho budeme řešit pro různé hodnoty penalizačního parametru a budeme sledovat vliv těchto hodnot na výsledky penalizačního problému.

1. Penalizační funkce

V této kapitole se budeme zabývat jedním konkrétním způsobem řešení nelineárních optimalizačních úloh a to právě převedením na úlohu hledání volného extrému, kdy jsou omezení převedena do účelové funkce pomocí takzvané penalizační funkce. Řešení tohoto problému dostaneme postupným zvyšováním penalizačního parametru, toto řešení limitně odpovídá řešení původní nelineární optimalizační úlohy. Kapitola vychází z [1][Chapter 9].

1.1 Základní problém

Zde si předvedeme základní problém, ze kterého vycházíme. Ukážeme si, co budeme požadovat po penalizační funkci a uvedeme si jednoduchý příklad pro demonstraci penalizační funkce. Pro motivaci použití penalizační funkce uvažujme následující problém:

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x}) = \mathbf{0}, \end{aligned} \tag{1.1}$$

kde $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})), \mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \ldots, h_l(\mathbf{x}))$, m, l $\in \mathbb{N}$, $\mathbf{x} \in \mathbb{R}^n$ a $\mathbf{0}$ je nulový vektor. Nechť μ je velké číslo. Nahraďme tento problém následujícím problémem :

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) + \mu\alpha(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in \mathbb{R}^n, \end{aligned} \tag{1.2}$$

kde $\alpha(\mathbf{x}) = \mathbf{h}^\top(\mathbf{x})\,\mathbf{h}(\mathbf{x}) + \text{maximum}\,\{0, \mathbf{g}(\mathbf{x})\}$. Triviálně pozorujeme, že optimální řešení výše uvedeného problému (1.2) musí mít $\alpha(\mathbf{x})$ blízko nule, jinak se projeví vliv velké penalizace skrz parametr μ. Povšimneme si, že pro omezení $\mathbf{g}(\mathbf{x})$ není možné použít penalizaci $\mathbf{g}^\top(\mathbf{x})\,\mathbf{g}(\mathbf{x})$, neboť zde dochází k penalizaci nezávisle na $\mathbf{g}(\mathbf{x}) > \mathbf{0}$, a nebo $\mathbf{g}(\mathbf{x}) < \mathbf{0}$. Tedy v tomto případě by nám penalizační funkce nefungovala v souladu s podmínkou $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$, je tedy nutné použít například maximum $\{0, \mathbf{g}(\mathbf{x})\}$, která penalizuje pro $\mathbf{g}(\mathbf{x}) > \mathbf{0}$ (maximum $\{0, \mathbf{g}(\mathbf{x})\} > 0$), ale pro $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ nedochází k žádné penalizaci (maximum $\{0, \mathbf{g}(\mathbf{x})\} = 0$).

Tedy po penalizační funkci vyžadujeme, aby pro body, co nesplňují podmínky(tj. jedná se o nepřípustné body), došlo k penalizaci a pro body, jež splňují podmínky, nedošlo k žádné penalizaci. Nyní zavedeme definici penalizační funkce, viz [1][str. 471].

Definice 1 (penalizační funkce)**.** *Nechť jsou dány omezení tvaru $g_i(\boldsymbol{x}) \leq 0$ a $h_j(\boldsymbol{x}) = 0$, pro $i = 1,\ldots,m$, $j = 1,\ldots,l$ a $\boldsymbol{x} \in \mathbb{R}^n$ pak funkci*

$$\alpha(\boldsymbol{x}) = \sum_{i=1}^{m} \phi[g_i(\boldsymbol{x})] + \sum_{j=1}^{l} \psi[h_j(\boldsymbol{x})] \tag{1.3}$$

nazýváme penalizační funkci *jsou-li ϕ a ψ spojité funkce splňující pro $\forall y \in \mathbb{R}^n$:*

$$\begin{array}{llcll} \phi(y) = 0 & \textit{když } y \leq 0 & \textit{a} & \phi(y) > 0 & \textit{když } y > 0 \\ \psi(y) = 0 & \textit{když } y = 0 & \textit{a} & \psi(y) > 0 & \textit{když } y \neq 0. \end{array}$$

Typicky jsou funkce ϕ a ψ tvaru:

$$\phi(y) = [max\{0, y\}]^p,$$
$$\psi(y) = |y|^q,$$

kde p, q $\in \mathbb{N}$. Tím pádem je penalizační funkce $\alpha(\mathbf{x})$ tvaru:

$$\alpha(\mathbf{x}) = \sum_{i=1}^{m}[max\{0, g_i(\mathbf{x})\}]^p + \sum_{j=1}^{l}[|h_j(\mathbf{x})|^q]. \tag{1.4}$$

.

Definice 2 (pomocná funkce)**.** *Funkci f(**x**) + $\mu\alpha$(**x**), kde α(**x**) je penalizační funkce, nazýváme* pomocnou funkcí.

V následujícím jednoduchém příkladu demonstrujeme převod optimalizační úlohy s vázaným extrémem (1.1) na úlohu hledaní volného extrému pomocné funkce (1.2).

Příklad 1. Uvažujme následující optimalizační úlohu s vázaným extrémem:

$$\begin{aligned} &\text{minimalizovat} && x^2 + y^2 \\ &\text{za podmínek} && x + y - 1 = 0. \end{aligned}$$

Abychom získali optimální řešení, tak z druhé rovnice vyjádříme x = 1 − y. Dosazením do účelové funkce a po následném zderivováním $(1 - y)^2 + y^2$, dostáváme rovnici $4y - 2 = 0$, neboť druhá derivace této funkce je kladná, jedná se o lokální minimum ([3] [Věta 5.4.14]), tím pádem nacházíme optimální řešení v bodě (1/2, 1/2). Optimální hodnota účelové funkce je 1/2. Nechť μ je velké číslo, uvažujme následující problém:

$$\begin{aligned} &\text{minimalizovat} && x^2 + y^2 + \mu(x + y - 1)^2 \\ &\text{za podmínek} && (x, y) \in \mathbb{R}^2. \end{aligned}$$

Penalizační funkce je zde tvaru $\alpha(x, y) = (x + y - 1)^2$. Povšimneme si, že penalizační funkce je zde dobře definovaná, neboť penalizuje právě tehdy, když x + y - 1 $\neq$ 0. Pro každé μ je pomocná funkce konvexní. Tím pádem nutnou a postačují podmínkou optimality je nulový gradient funkce $F(x, y) := x^2+y^2+\mu(x+y-1)^2$. Tedy:

$$\frac{\partial F}{\partial x} = x + \mu(x + y - 1) = 0,$$
$$\frac{\partial F}{\partial y} = y + \mu(x + y - 1) = 0.$$

Z první rovnice vyjádříme $x = \dfrac{\mu - \mu y}{1 + \mu}$, dosazením do druhé rovnice dostáváme:

$$y = \frac{\mu}{1 + 2\mu} \qquad x = \frac{\mu}{1 + 2\mu} \tag{1.5}$$

Tedy s optimálním řešením pomocné funkce se můžeme libovolně přiblížit k řešení originálního problému, pokud zvolíme μ dostatečně velké.

1.2 Penalizační metoda vnějšího bodu

V této části uvedeme Lemma 1 a Větu 2, která nás opravňuje převádět optimalizační úlohu s podmínkami typu rovností a nerovností (1.1), na problém hledání volného extrému s penalizační funkcí (1.2). Uvažujme dva následující obecné problémy:

Primární problém(PRP):

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & g_i(\mathbf{x}) \leq 0 \text{ pro i} = 1, \ldots, \text{m}, \\ & h_j(\mathbf{x}) = 0 \text{ pro j} = 1, \ldots, \text{l}, \\ & \mathbf{x} \in X, \end{aligned}$$

kde f, $g_1, \ldots, g_m, h_1, \ldots, h_l$ jsou spojité funkce na $\mathbb{R}^n$ a $X \subset \mathbb{R}^n$ je neprázdná množina, která je typicky tvaru X = $\{\mathbf{d} \leq \mathbf{x} \leq \mathbf{c}\}$, kde $\mathbf{d} \leq \mathbf{c}$ jsou vektory konstant. (PRP) je vlastně reformulace úvodního hlavního problému nelineárního programování (1.1).

Penalizační problém(PEP):
Nechť α je spojitá *penalizační funkce*, pak pro $\mu \geq 0$ definujeme *Penalizační problém* jako:

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) + \mu\alpha(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in X. \end{aligned}$$

Penalizační přístup se pomocí parametru $\mu \geq 0$ snaží najít nejmenší hodnotu pomocné funkce f($\mathbf{x}$) + $\mu\alpha(\mathbf{x})$ a z těchto minim vybere μ, pro které je funkce největší, tedy formálně:

$$\begin{gathered} \text{Sup } \theta(\mu) \\ \text{za podmínek } \mu \geq 0 \end{gathered}$$

kde $\theta(\mu) := \inf \{f(\mathbf{x}) + \mu\alpha(\mathbf{x}) : \mathbf{x} \in X\}$.

Náš cíl v této sekci je ukázat, že platí následující rovnosti:

$$\inf\{f(\mathbf{x}) : \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} = \sup_{\mu \geq 0} \theta(\mu) = \lim_{\mu \to \infty} \theta(\mu). \tag{1.6}$$

Kde $\mathbf{g}(\mathbf{x})$ a $\mathbf{h}(\mathbf{x})$ jsou definované jako v 1.1. Tento výsledek nás opravňuje převádět Primární problém (PRP) na Penalizační problém (PEP), neboť zvolením dostatečné velkého μ v (PEP) se můžeme libovolně přiblížit k optimálnímu řešení (PRP). Pro důkaz těchto rovností je zapotřebí následující lemma.

Lemma 1. *Nechť f, $g_1, \ldots, g_m, h_1, \ldots, h_l$ jsou spojité funkce na $\mathbb{R}^n$ a $X \subset \mathbb{R}^n$ je neprázdná množina, dále předpokládejme, že α je spojitá penalizační funkce na $\mathbb{R}^n$ a pro každé μ existuje $\boldsymbol{x}_\mu \in X$ takové, že $\theta(\mu) = f(\boldsymbol{x}_\mu) + \mu\alpha(\boldsymbol{x}_\mu)$. Pak platí následující tvrzení:*

1.

$$\inf\{f(\boldsymbol{x}) : \boldsymbol{x} \in X,\ \boldsymbol{g(x)} \leq \boldsymbol{0}, \boldsymbol{h(x)} = \boldsymbol{0}\} \geq \sup_{\mu \geq 0} \theta(\mu)$$

2. $f(\boldsymbol{x}_\mu)$ je neklesající funkce $\mu \geq 0$, $\theta(\mu)$ je neklesající funkce μ a $\alpha(\boldsymbol{x}_\mu)$ je nerostoucí funkce μ.

Důkaz. Viz [1][Lemma 9.2.1]

□

Věta 2. *Uvažujme Primární problém (PRP). Předpokládejme, že tento problém má přípustné řešení a nechť α je spojitá penalizační funkce. Dále předpokládejme, že pro každé μ existuje řešení $\boldsymbol{x}_\mu \in X$ Penalizačního problému (PEP) a také, že posloupnost $\{\boldsymbol{x}_\mu\}$ je obsažena v kompaktní podmnožině X. Pak platí:*

$$\inf\{f(\boldsymbol{x}) : \boldsymbol{x} \in X, \boldsymbol{g(x)} \leq \boldsymbol{0}, \boldsymbol{h(x)} = \boldsymbol{0}\} = \sup_{\mu \geq 0} \theta(\mu) = \lim_{\mu \to \infty} \theta(\mu)$$

a také, že limita $\overline{\boldsymbol{x}}$ libovolné podposloupnosti $\{\boldsymbol{x}_\mu\}$ je optimální řešení Primárního problému (PRP) a $\mu\alpha(\boldsymbol{x}_\mu) \to 0$ pro $\mu \to \infty$.

Důkaz. Provedeme pouze náznak důkazu, kompletní důkaz lze nalézt v [1] [Theorem 9.2.2].
Neboť dle 2.části Lemma 1 je funkce $\theta(\mu)$ neklesající v μ, pak má limitu a platí

$$\sup_{\mu \geq 0} \theta(\mu) = \lim_{\mu \to \infty} \theta(\mu),$$

dále je třeba dokázat, že $\alpha(\mathbf{x}_\mu) \to 0$ pro $\mu \to \infty$. To se ukáže sporem pomocí Lemma 1. Nechť $\{\mathbf{x}_{\mu_k}\}$ je libovolná konvergentní podposloupnost $\{\mathbf{x}_\mu\}$ a nechť $\overline{\mathbf{x}}$ je jeho limita, neboť platí:

$$\sup_{\mu \geq 0} \theta(\mu) \geq \theta(\mu_k) = f(\mathbf{x}_{\mu_k}) + \mu\alpha(\mathbf{x}_{\mu_k}) \geq f(\mathbf{x}_{\mu_k})$$

a zároveň f je spojitá a $\{\mathbf{x}_{\mu_k}\} \to \overline{\mathbf{x}}$ pak:

$$\sup_{\mu \geq 0} \theta(\mu) \geq f(\overline{\mathbf{x}}). \tag{1.7}$$

Tedy víme, že $\alpha(\mathbf{x}_\mu) \to 0$ pro $\mu \to \infty$ a $\alpha(\overline{\mathbf{x}}) = 0$ z toho plyne, že $\overline{\mathbf{x}}$ je přípustné řešení (PRP). Dále z nerovnosti 1.7 a z 1.části Lemma 1 nutně vyplývá, že $\overline{\mathbf{x}}$ je optimální řešení (PRP) a tedy také $\sup_{\mu \geq 0} \theta(\mu) = f(\overline{\mathbf{x}})$. Zbývá dokázat, že $\mu\alpha(\mathbf{x}_\mu) \to 0$ pro $\mu \to \infty$. Přepíšeme $\mu\alpha(\mathbf{x}_\mu) = \theta(\mu) - f(\mathbf{x}_\mu)$. Oba dva výrazy napravo konvergují k $f(\overline{\mathbf{x}})$ pro $\mu \to \infty$, tedy $\mu\alpha(\mathbf{x}_\mu)$ konverguje k 0 pro $\mu \to \infty$.

□

Kompaktnost podmnožiny X je nutným předpokladem Věty 2. Pokud by daná podmnožina X nebyla kompaktní, pak by mohl nastat případ, že (PEP) a (PRP) nebyly ekvivalentní problémy dle Věty 2. Nicméně v praktických úlohách s předpokladem kompaktnosti nebývá problém, neboť množina X je většinou tvaru:

$$\begin{aligned} d_1 \leq x_1 \leq c_1, \ldots, d_n \leq x_n \leq c_n, \\ d_1 \leq c_1, \ldots, d_n \leq c_n, \\ c_1, \ldots, c_n \in \mathbb{R}\ d_1, \ldots, d_n \in \mathbb{R}. \end{aligned}$$

Z Věty 2 plyne, že optimální řešení $\mathbf{x}_\mu$ (PEP) konvergují k množině přípustných řešení. Jestliže zvolíme dostatečně velké μ, pak s hodnotou $f(\mathbf{x}_\mu) + \mu\alpha(\mathbf{x}_\mu)$ se lze libovolně přiblížit k optimální hodnotě účelové funkce (PRP).

Jeden z možných přístupu řešení Penalizačního problému (PEP) je vyřešit posloupnost problémů:

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) + \mu_k\alpha(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in \mathbb{R}^n, \end{aligned}$$

pro rostoucí parametr μ_k. Posloupnost optimálních bodů $\{\mathbf{x}_{\mu_k}\}$ typicky nesplňuje vazební podmínky(tj. jedná se o nepřípustné body), avšak s postupným zvětšováním penalizačního parametru μ se posloupnost optimálních bodů blíží k přípustné množině a také k optimálnímu řešení Primárního problému (PRP), jak později uvidíme na příkladu. Proto označení *Metoda vnějšího bodu*.

1.3 Podmínky optimality, Lagrangeovy multiplikátory

Nejprve si uvedeme definice *regulárního řešení* a *KKT bodu*, to je bod, který splňuje KKT podmínky(tzn. Karushovy-Kuhnovy-Tuckerovy podmínky), viz [1][Theorem 4.3.7].

Definice 3 (KKT bod, regulární řešení)**.** *Nechť $X \in \mathbb{R}^n$ je neprázdná otevřená množina a f: $\mathbb{R}^n \to \mathbb{R}$, g_i: $\mathbb{R}^n \to \mathbb{R}$ $i = 1, \ldots, m$, $h_j : \mathbb{R}^n \to \mathbb{R}$ $j = 1, \ldots, l$. Uvažujme penalizační problém (PRP). Nechť $\overline{\boldsymbol{x}}$ je přípustné řešení a definujme $I := \{i : g_i(\overline{\boldsymbol{x}}) = 0\}$. Dále předpokládejme, že f a g_i jsou diferencovatelné v bodě $\overline{\boldsymbol{x}}$ pro $i = 1, \ldots, m$, h_j jsou spojitě diferencovatelné v bodě $\overline{\boldsymbol{x}}$ pro $j = 1, \ldots, l$ a platí, že $\nabla g_i(\overline{\boldsymbol{x}})$ a $\nabla h_j(\overline{\boldsymbol{x}})$ jsou lineárně nezávislé pro $i \in I$, $j = 1, \ldots, l$, pak bod $\overline{\boldsymbol{x}}$ nazveme* regulární řešení.
Jestliže $\overline{\boldsymbol{x}}$ splňuje následující KKT podmínky:

$$\begin{aligned} \nabla f(\overline{\boldsymbol{x}}) + \sum_{i=1}^{m} u_i \nabla g_i(\overline{\boldsymbol{x}}) + \sum_{j=1}^{l} v_j \nabla h_j(\overline{\boldsymbol{x}}) = 0 & \\ u_i g_i(\overline{\boldsymbol{x}}) = 0 \qquad & i = 1, \ldots, m, \\ u_i \geq 0 \qquad & i = 1, \ldots, m, \end{aligned}$$

pak bod $\overline{\boldsymbol{x}}$ nazveme KKT bod *s příslušnými Lagrangeovými multiplikátory (u_i,v_j).*

Za určitých předpokladů je možné využít posloupnost $\{\mathbf{x}_\mu\}$ pro získání KKT bodu s příslušnými Lagrangeovými multiplikátory. Nechť $X = \mathbb{R}^n$ a uvažujme (PRP). Nechť α je penalizační funkce a navíc platí, že ψ a ϕ jsou spojitě diferencovatelné funkce, splňující $\phi'(\mathrm{y}) \geq 0$ pro všechny y a $\phi'(\mathrm{y}) = 0$ pro všechny y ≤ 0. Dále předpokládejme, že platí předpoklady Věty 2. Označme $\mathbf{x}_\mu$ jako optimální řešení(z Věty 2), pak musí platit, že gradient funkce f($\mathbf{x}$) + $\mu\alpha(\mathbf{x})$ se v bodě $\mathbf{x}_\mu$ rovná 0. Tedy:

$$\nabla f(\mathbf{x}_\mu) + \sum_{i=1}^{m} \mu\phi'[g_i(\mathbf{x}_\mu)]\nabla g_i(\mathbf{x}_\mu) + \sum_{j=1}^{l} \mu\psi'[h_j(\mathbf{x}_\mu)]\nabla h_i(\mathbf{x}_\mu) = 0 \qquad \forall\mu. \tag{1.8}$$

Nechť $\overline{\mathbf{x}}$ je limitou posloupnosti $\{\mathbf{x}_\mu\}$. Označme I := $\{i : g_i(\overline{\mathbf{x}}) = 0\}$, tedy platí $g_i(\overline{\mathbf{x}}) < 0$ pro $\forall$ i $\notin$ I a tím pádem z Věty 2 plyne, že pro μ dostatečně velké je i $g_i(\mathbf{x}_\mu) < 0$, to společně se zesilněnými předpoklady na ϕ implikuje $\mu\phi'[g_i(\mathbf{x}_\mu)]$ $= 0$ pro $\forall$ i $\notin$ I. Tedy můžeme přepsat rovnost (1.8) na:

$$\nabla f(\mathbf{x}_\mu) + \sum_{i\in I}(\mathbf{u}_\mu)_i \nabla g_i(\mathbf{x}_\mu) + \sum_{j=1}^{l}(\mathbf{v}_\mu)_j \nabla h_j(\mathbf{x}_\mu) = 0 \qquad \text{pro } \mu \text{ dostatečně velké,} \tag{1.9}$$

kde $\mathbf{u}_\mu$ a $\mathbf{v}_\mu$ jsou vektory tvaru:

$$\begin{aligned} (\mathbf{u}_\mu)_i &= \mu\phi'[g_i(\mathbf{x}_\mu)] \geq\ 0 \qquad && \forall i \in I, \\ (\mathbf{v}_\mu)_j &= \mu\psi'[h_j(\mathbf{x}_\mu)] && j = 1, \ldots, l. \end{aligned} \tag{1.10}$$

Nechť $\overline{\mathbf{x}}$ je regulární řešení. Potom víme, že existují jednoznačné Lagrangeovy multiplikátory $\overline{u_i} \geq 0$ pro $i \in I$ a $\overline{v_j}$ pro $j = 1, \ldots, l$ takové, že platí:

$$\nabla f(\overline{\mathbf{x}}) + \sum_{i\in I} \overline{u_i} \nabla g_i(\overline{\mathbf{x}}) + \sum_{j=1}^{l} \overline{v_j} \nabla h_j(\overline{\mathbf{x}}) = 0. \tag{1.11}$$

Neboť g_i, h_j, ϕ_i, ψ_j jsou spojitě diferencovatelné funkce, pro všechna $i \in I$, $j = 1, \ldots, l$ a navíc platí, že $\{\mathbf{x}_\mu\} \to \overline{\mathbf{x}}$ a jelikož $\overline{\mathbf{x}}$ je regulární řešení, pak musí platit:

$$\begin{aligned} (\mathbf{u}_\mu)_i &\to \overline{u_i} \qquad && \forall i \in I, \\ (\mathbf{v}_\mu)_j &\to \overline{v_j} && j = 1, \ldots, l. \end{aligned}$$

Tím pádem pro dostatečně velké μ můžeme KKT bod s příslušnými Lagrangeovými multiplikátory aproximovat multiplikátory definovanými v (1.10).

Uvažujme Primární problém (PRP) a penalizační funkci tvaru:

$$\alpha(\mathbf{x}) = \sum_{i=1}^{m}[max\{0, g_i(\mathbf{x})\}]^p + \sum_{j=1}^{l}[h_j(\mathbf{x})]^q,$$

kde $p, q \in \mathbb{N}$.

$$\begin{aligned} \phi(y) &= [max\{0, y\}]^p \qquad && \phi'(y) = p[max\{0, y\}]^{p-1}, \\ \psi(y) &= y^q && \psi'(y) = qy^{q-1}. \end{aligned} \tag{1.12}$$

Pak z (1.10) dostáváme:

$$\begin{aligned} (\mathbf{u}_\mu)_i &= \mu p[max\{0, y\}]^{p-1} \qquad && \forall i \in I, \\ (\mathbf{v}_\mu)_j &= \mu q y^{q-1} && j = 1, \ldots, l. \end{aligned} \tag{1.13}$$

Příklad 2. Vraťme se k *Příkladu* 1. Spočtěme Lagrangeův multiplikátor λ.
$L(x, y, \lambda) := x^2 + y^2 + \lambda(x + y - 1)$:

$$\frac{\partial L}{\partial x} = 2x + \lambda = 0 \qquad \text{(i)}$$
$$\frac{\partial L}{\partial y} = 2y + \lambda = 0 \qquad \text{(ii)}$$
$$\frac{\partial L}{\partial \lambda} = x + y - 1 = 0 \qquad \text{(iii)}$$

z (i) a (ii) rovnice vyjádříme $x = y = -\lambda/2$ a dosazením do (iii) získáme $\lambda = -1$
Pomocí výsledku (1.5) vyjádříme:

$$\mathbf{x}_\mu = [\frac{\mu}{1+2\mu}, \frac{\mu}{1+2\mu}] \qquad h(\mathbf{x}_\mu) = \frac{2\mu}{1+2\mu} - 1 = \frac{-1}{1+2\mu},$$
$$\text{a tedy z (1.13) pro q = 2} \qquad \mathbf{v}_\mu = \frac{-2\mu}{1+2\mu}.$$

Spočítáme $\lim_{\mu \to \infty} \mathbf{v}_\mu = -1$. Tedy $\mathbf{v}_\mu \to \lambda$.

1.4 Algoritmus řešení Penalizačního problému

Jeden z možných přístupu k řešení Penalizačního problému (PEP), je řešení posloupnosti (PEP) s postupně rostoucím μ, tedy s každou novou penalizační hodnotou μ optimalizujeme aktuální (PEP) s optimálním řešením z předchozího kroku.
Uvažujme Primární problém (PRP):

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x}) = \mathbf{0}, \\ & \mathbf{x} \in X, \end{aligned}$$

nechť $\alpha(x)$ je penalizační funkce.

Algoritmus
Inicializace: Nechť je dána přesnost $\epsilon > 0$. Zvolme počáteční bod $\mathbf{x}_1$, penalizační parametr $\mu_1 > 0$ a intenzitu $\beta > 1$. Nechť k := 1

1. Pomocí $\mathbf{x}_k$, vyřešme Penalizační problém (PEP):

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) + \mu_k \alpha(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in X. \end{aligned}$$

   Nechť $\mathbf{x}_{k+1}$ je optimální řešení.

2. Pokud $\mu_k \alpha(\mathbf{x}_{k+1}) < \epsilon$, pak jsme nalezli optimální řešení Primárního problému (PRP), jinak $\mu_{k+1} = \beta \mu_k$, $k \to k + 1$ a přejdeme na krok 1.

Iterace	μ_k	$\mathbf{x}_{k+1} = \mathbf{x}_{\mu_k}$	$f(\mathbf{x}_{k+1})$	$\alpha(\mathbf{x}_{k+1})$	$\mu_k\alpha(\mathbf{x}_{k+1})$
1	0,1	(1,1801; 1,9020)	0,0421	0,5878	0,0588
2	1	(1,1891; 1,6235)	0,1775	0,1920	0,1920
3	10	(1,0658; 1,1792)	0,6780	0,0139	0,1385
4	100	(1,0095; 1,0239)	0,9528	0,0002	0,0231
5	1000	(1,0010; 1,0025)	0,9950	0,0000	0,0025

Tabulka 1.1: Použití algoritmu pro řešení Penalizačního problému.

Výše uvedený algoritmus je efektivně použitelný pouze, pokud umíme efektivně vyřešit 1.krok.

Příklad 3. Uvažujme následující problém:

$$\begin{aligned} \text{minimalizovat} \quad & (x-1)^2 + (y-2)^2 \\ \text{za podmínek} \quad & x^3 - y = 0, \\ & x - y = 0. \end{aligned} \tag{1.14}$$

Pro vyřešení výše uvedeného problému uvažujme pro k $\in \mathbb{N}$ pomocnou funkci:

$$(x-1)^2 + (y-2)^2 + \mu_k((x^3-y)^2 + (x-y)^2).$$

Řešení problému (1.14) je zřejmě bod (1,1). Pro použití algoritmu (ALG) definujme $\mu_1 := 0{,}1$, $\beta := 10$, $\epsilon := 0{,}01$ a počáteční bod $\mathbf{x}_1$ zvolíme tak, aby platilo, že hodnota účelové funkce je v $\mathbf{x}_1$ rovna 0. Tedy $\mathbf{x}_1 := (1{,}2)$. Tabulka 1.1 ukazuje výsledky pro jednotlivá k.Výpočet skončil pro k = 5, neboť $0{,}0025 = \mu_5\alpha(\mathbf{x}_6) < \epsilon = 0{,}01$.

Povšimne si, že výsledky v Tabulce 1.1 odpovídají tvrzení Lemma 1 a Věty 2. Funkce $f(\mathbf{x}_{k+1})$ je neklesající a funkce $\alpha(\mathbf{x}_{k+1})$ je nerostoucí a pro zvětšující se k konverguje k 0. Zároveň je patrné, že s rostoucím k konverguje $\mathbf{x}_{\mu_k}$ k (1,1).

2. Exaktní penalizační funkce

Dosud jsme při řešení penalizačního problému byli nuceni přibližovat se s parametrem μ k nekonečnu. Cílem této kapitoly je představit speciální případy penalizačních funkcí, které nevyžadují limitní přiblížení μ k nekonečnu. Jinými slovy řeší Penalizační problém (PEP) pro rozumné konečné μ. Tyto penalizační funkce nazýváme Exaktními penalizačními funkcemi.

Uvažujme primární problém (PRP) a k němu příslušnou *Absolutní penalizační funkci*, která je tvaru 1.4 pro $p = q = 1$, tedy:

$$\alpha(\mathbf{x}) = \sum_{i=1}^{m} max\{0, g_i(\mathbf{x})\} + \sum_{j=1}^{l} |h_j(\mathbf{x})| \tag{2.1}$$

Dle Věty 3, uvedené níže, je *Absolutní penalizační funkce* tvaru 2.1 exaktní. Definujme pomocnou funkci F_E s penalizační funkcí 2.1 jako:

$$F_E(\mathbf{x}) = f(\mathbf{x}) + \mu[\sum_{i=1}^{m} max\{0, g_i(\mathbf{x})\} + \sum_{j=1}^{l} |h_j(\mathbf{x})|] \tag{2.2}$$

Pokud zesílíme předpoklady na funkce f, g_i a h_j, pak nám následující Věta 3 dává existenci konečného parametru μ, díky němuž získáme optimální hodnotu (PRP) pomocí minimalizace pomocné funkce F_E.

Věta 3. *Uvažujme primární problém (PRP):*

$$\begin{array}{rll} \textit{minimalizovat} & f(\boldsymbol{x}) & \\ \textit{za podmínek} & g_i(\boldsymbol{x}) \leq 0 & i = 1, \ldots, m, \\ & h_j(\boldsymbol{x}) = 0 & j = 1, \ldots, l. \end{array}$$

Nechť $\overline{\mathbf{x}}$ *je KKT bod s příslušnými Lagrangeovými multiplikátory* $\overline{u_i} \geq 0$ *pro* $i \in I$ *a* $\overline{v_j}$ *pro* $j = 1, \ldots, l$ *s příslušnými omezeními tvaru rovnosti a nerovnosti, respektive s* $I := \{i \in \{1, \ldots, m\} : g_i(\overline{\mathbf{x}}) = 0\}$. *Dále předpokládejme, že* f *a* g_i, $i \in I$ *jsou konvexní funkce a* h_j, $j = 1, \ldots, l$ *jsou afinní funkce. Pak pro* $\mu \geq max\{\overline{u_i},\ i \in I,\ |\overline{v_j}|,\ j = 1, \ldots, l\}$ *bod* $\overline{\mathbf{x}}$ *minimalizuje pomocnou funkci* F_E *s* Absolutní penalizační funkci *definovanou v 2.2.*

Důkaz. Viz [1][Theorem 9.3.1]

□

Příklad 4. Opět uvažujme *Příklad* 1. V *Příkladu* 2 jsme spočítali Lagrangeův multiplikátor $\lambda = -1$ příslušný k KKT bodu minima $\overline{\mathbf{x}} = (1/2,\ 1/2)$. Označme $\overline{v} = \lambda = -1$. Pak funkce F_E definovaná v 2.2 je tvaru:

$$F_E = x^2 + y^2 + \mu\,|x + y - 1|\,.$$

Pro $\mu = 0$ je F_E minimalizována v bodě (0, 0). Pro $\mu > 0$ je minimalizace F_E ekvivalentní úloze:

$$\begin{array}{rl} \text{minimalizovat} & x^2 + y^2 + \mu z \\ \text{za podmínek} & z \geq x + y - 1, \\ & z \geq -x - y + 1. \end{array}$$

Lagrangeova funkce je tvaru:

$$L := x^2 + y^2 + \mu^+(x + y - 1) + \mu^-(-x - y + 1),$$

tedy KKT podmínky vyžadují aby platilo následující:

$$\begin{aligned} 2x + (\mu^+ - \mu^-) &= 0 \qquad & 2y + (\mu^+ - \mu^-) &= 0, \\ \mu^+(z - x - y + 1) &= 0 \qquad & \mu^-(z + x + y - 1) &= 0, \end{aligned}$$

kde

$$\mu = \mu^+ + \mu^- \qquad \mu^+, \mu^- \geq 0 \qquad z = |x + y - 1|,$$

příklad rozdělíme na 3 části:

- $\mathbf{x + y < 1}$: $z = -x - y + 1$ a $\mu^+ = 0$ tím pádem $\mu^- = \mu$ a $x = y = \mu/2$. Našli jsme tedy KKT bod $[(\mu/2, \mu/2), \mu]$, kde $0 \leq \mu < 1$.
- $\mathbf{x + y = 1}$: $z = 0$ a platí $\mu^+(-x-y+1) = \mu^-(x+y-1) = 0$ vyřešením této soustavy dostáváme, že $x = y = 1/2$ a $1/2 = x = -(\mu^+ - \mu^-)/2$. Vyřešíme:

  $$\begin{aligned} 1/2 &= -(\mu - 2\mu^-)/2 \qquad & 1/2 &= -(-\mu + 2\mu^+)/2, \\ \mu^- &= (\mu + 1)/2 \qquad & \mu^+ &= (\mu - 1)/2, \end{aligned}$$

  našli jsme tedy KKT bod $[(1/2, 1/2), (\mu - 1)/2, (\mu + 1)/2]$, kde $\mu \geq 1$
- $\mathbf{x + y > 1}$: $z = x + y - 1$ a $\mu^- = 0$, pak tedy $2x = -\mu^+ = 2y$ přepíšeme $x = y = -\mu^+/2$, dále $\mu = \mu^+$ a tím pádem $x + y = -\mu^+/2 - \mu^+/2 = -\mu$ což je spor s $\mu \geq 0$.

Dohromady máme, že pro $0 \leq \mu < 1$ je F_E minimalizována v bodě $(\mu/2, \mu/2)$ a pro $\mu \geq 1$ zůstane F_E minimální v bodě $(1/2, 1/2)$, tedy v optimálním bodě *Příkladu* 1.

2.1 Celočíselné nelineární programování

V této části bude naším úkolem diskutovat řešení nelineárních úloh s binárními proměnnými. Podkapitola je zpracována dle článku [2]. Nejprve si zadefinujme dva problémy: *Binární problém* a k němu příslušný *Binární penalizační problém.* Ekvivalenci těchto dvou problému budeme dokazovat ve Větě 5.

Binární problém (BP)

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in T \cap \{0, 1\}^n, \end{aligned}$$

kde $T \subseteq \mathbb{R}^n$ a f je funkce splňující podmínku ($\boldsymbol{A}$) z Věty 4:

Binární penalizační problém (BPP)

$$\begin{aligned} \text{minimalizovat} \quad & f(\mathbf{x}) + \varphi(\mathbf{x}, \varepsilon) \\ \text{za podmínek} \quad & \mathbf{x} \in T, 0 \leq \mathbf{x} \leq e, \end{aligned}$$

kde $\varphi(\mathbf{x}, \varepsilon)$ je vhodná penalizační funkce a $\varepsilon > 0$ penalizační parametr, který je dostatečně malý(tzn. sebemenší porušení binarity je penalizováno velkou hodnotou). Tedy namísto původního celočíselného problému, dostaneme složitý nelineární problém. Jak uvidíme dále u příkladů penalizačních funkcí 2.3 - 2.7, tak vhodná penalizační funkce musí penalizovat porušení binarity $\mathbf{x}$.

Binární problém (BP) je speciálním případem našeho úvodního primárního problému (PRP) s množinou přípustných řešení: $T \cap \{0,1\}^n$. Penalizační funkce $\varphi(\mathbf{x}, \varepsilon)$ je trochu jiného tvaru, než na jaký jsme byli dosud zvyklí z definice *Penalizační funkce* 1 . Zde je penalizační parametr $\varepsilon > 0$ o proti $\mu \geq 0$ a je součástí samotné penalizační funkce například:

$$\varphi(\mathbf{x}, \varepsilon) = \sum_{i=1}^{n} \log(x_i + \varepsilon) + \log((1 - x_i) + \varepsilon), \tag{2.3}$$

$$\varphi(\mathbf{x}, \varepsilon) = \sum_{i=1}^{n} -(x_i + \varepsilon)^{-p} - (1 - x_i + \varepsilon)^{-p}, \tag{2.4}$$

$$\varphi(\mathbf{x}, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} 1 - \exp(-\alpha x_i) + 1 - \exp(-\alpha(1 - x_i)), \tag{2.5}$$

$$\varphi(\mathbf{x}, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} (x_i + \varepsilon)^q + (1 - x_i + \varepsilon)^q, \tag{2.6}$$

$$\varphi(\mathbf{x}, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} (1 + \exp(-\alpha x_i))^{-1} + (1 + \exp(-\alpha(1 - x_i)))^{-1}, \tag{2.7}$$

kde $\varepsilon, \alpha, p > 0$ a $0 < q < 1$, viz [2][str. 483].

K důkazu ekvivalence binárního problému (BP) a binárního penalizačního problému (BPP) si zadefinujeme zobecnění těchto dvou problému, jejichž ekvivalenci ukazuje Věta 4.
Uvažujme problém s vázaným extrémem:

$$\min_{x \in W} f(\mathbf{x}), \tag{2.8}$$

kde $W \subset \mathbb{R}^n$ a $f : \mathbb{R}^n \to \mathbb{R}$.
Pro $\varepsilon \in (0, \infty)$, uvažujme následující problém:

$$\min_{x \in X} f(\mathbf{x}) + \varphi(\mathbf{x}, \varepsilon), \tag{2.9}$$

kde $W \subseteq X \subset \mathbb{R}^n$ a $\varphi : \mathbb{R}^n \to \mathbb{R}$ jako funkce proměnné $\mathbf{x}$ pro pevné ε.

Povšimne si podobnosti 2.8 a primárního problému (PRP). Zde nemáme přímo definovaná omezení tvaru rovností a nerovností, ale máme již obecnou množinu přípustných řešení W, dokonce nevyžadujeme ani spojitost f, avšak jak uvidíme ve Větě 4, další předpoklady na f a φ jsou nutné pro dokázaní ekvivalence problémů 2.8 a 2.9.

Věta 4. *Nechť W a X jsou kompaktní množiny. Nechť $\|\cdot\|$ je vhodně zvolená norma. Nechť platí:*

(A) *f je omezená na X a existuje otevřená množina $A \supset W$ a reálné číslo $L > 0$ takové, že pro všechna $\boldsymbol{x}, \boldsymbol{y} \in A$, f splňuje podmínku:*

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq L\|\boldsymbol{x} - \boldsymbol{y}\|. \tag{2.10}$$

Dále nechť φ splňuje následující podmínky:

(B) *Pro všechna $\boldsymbol{x}, \boldsymbol{y} \in W$ a pro všechny $\varepsilon \in (0, \infty)$*

$$\varphi(\boldsymbol{x}, \varepsilon) = \varphi(\boldsymbol{y}, \varepsilon). \tag{2.11}$$

(C) *Existuje $\hat{\varepsilon} \in (0, \infty)$, $\forall \boldsymbol{z} \in W$ existuje okolí $S(\boldsymbol{z})$ takové, že $\forall \boldsymbol{x} \in S(\boldsymbol{z}) \cap (X \backslash W)$ a $\varepsilon \in (0, \hat{\varepsilon}]$ platí:*

$$\varphi(\boldsymbol{x}, \varepsilon) - \varphi(\boldsymbol{z}, \varepsilon) \geq \hat{L} \|\boldsymbol{x} - \boldsymbol{z}\|, \tag{2.12}$$

kde $\hat{L} > L$. Dále nechť $S = \cup_{z \in W} S(z)$ a existuje $\overline{\boldsymbol{x}} \notin S$ takové, že :

$$\lim_{\varepsilon \to 0^+} [\varphi(\overline{\boldsymbol{x}}, \varepsilon) - \varphi(\boldsymbol{z}, \varepsilon)] = \infty \qquad \forall z \in W, \tag{2.13}$$

$$\varphi(\boldsymbol{x}, \varepsilon) \geq \varphi(\overline{\boldsymbol{x}}, \varepsilon) \qquad \forall x \in X \backslash S, \forall \varepsilon > 0. \tag{2.14}$$

Pak existuje $\tilde{\varepsilon} \in \mathbb{R}$ takové, že $\forall \varepsilon \in (0, \tilde{\varepsilon}]$ problémy 2.8 a 2.9 mají stejná minima.

Důkaz. Viz [2][Theorem 2.1]

□

Předpoklad **(B)** znamená, že všechna celočíselná řešení jsou penalizována stejně. Dále nerovnost 2.12 říká, že pro libovolné $\mathbf{x}$ co náleží do okolí přípustného bodu $\mathbf{z}$, platí, že rozdíl penalizaci je větší než $\hat{L}$-krát norma jejich rozdílu.

Následuje Věta 5, která vychází z Věty 4, jež ukazuje ekvivalenci naších dvou úvodních celočíselných problému: Binárního problému (BP) a Binárního penalizačního problému (BPP), pro všechny penalizační funkce 2.3 - 2.7.

Věta 5. *Položme*

$$W = \{\boldsymbol{x} \in T : \boldsymbol{x} \in \{0,1\}^n\}, \qquad X = \{\boldsymbol{x} \in T : 0 \leq \boldsymbol{x} \leq e\}.$$

Pak pro každou penalizační funkci 2.3 - 2.7 existuje $\overline{\varepsilon} > 0$ takové, že pro $\forall \varepsilon \in (0, \overline{\varepsilon}]$ problémy (BP) a (BPP) mají stejná minima.

Důkaz. Ekvivalenci problému (BP) a (BPP) dokážeme pomocí Věty 4. Neboť předpokládáme platnost **(A)** pro funkci *f*, stáčí ukázat, že každá penalizační funkce 2.3 - 2.7 splňuje **(B)** a **(C)** z Věty 4. Provedeme důkaz pro penalizační funkci 2.4, pro ostatní penalizační funkce je důkaz obdobný. Ve [2][Proposition 3.1] je uveden důkaz pro 2.3, my budeme postupovat obdobně. Uvažujme tedy penalizační funkci 2.4:

$$\varphi(\mathbf{x}, \varepsilon) = \sum_{i=1}^{n} -(x_i + \varepsilon)^{-p} - (1 - x_i + \varepsilon)^{-p},$$

kde $p > 0$.
Pro všechny $\mathbf{x} \in \{0,1\}^n$ platí:

$$\varphi(\mathbf{x}, \varepsilon) = n[-(\varepsilon)^{-p} - (1 + \varepsilon)^{-p}],$$

tím pádem platí předpoklad **(B)**.
Nyní se budeme zabývat chováním i-té funkce $\varphi_i(x_i, \varepsilon)$ na okolí přípustného bodu z_i. Chceme tedy ukázat platnost 2.12. Uvažujme libovolné $\mathbf{x}$ splňující: $\mathbf{x} \in S(\mathbf{z}) \cap (X \backslash W)$. Tím pádem nastávají tři různé případy:

1. $z_i = 0$ a $0 < x_i < \gamma$: φ_i splňuje na $[0, x_i]$ Lagrangeovu větu o střední hodnotě ([3][Věta 5.2.4]) a tedy platí:

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = p[(\tilde{x}_i + \varepsilon)^{-p-1} - (1 - \tilde{x}_i + \varepsilon)^{-p-1}]|x_i - z_i|,$$

kde $\tilde{x}_i \in (0, x_i)$. Zvolme $\gamma < 1/2$ a neboť $p > 0$ pak platí:

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq p[(\gamma + \varepsilon)^{-p-1} - (1 - \gamma + \varepsilon)^{-p-1}]|x_i - z_i|.$$

Odhadneme druhý člen:

$$(1 - \gamma + \varepsilon)^{-p-1} \leq (1 - 1/2 + \varepsilon)^{-p-1} = (1/2 + \varepsilon)^{-p-1} \leq (1/2)^{-p-1},$$

a tedy dostaváme:

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq p[(\gamma + \varepsilon)^{-p-1} - (1/2)^{-p-1}]|x_i - z_i|.$$

Zvolíme ε, γ tak aby platilo:

$$(\gamma + \varepsilon)^{-p-1} \geq \frac{\tilde{L}}{p} + (1/2)^{-p-1}, \tag{2.15}$$

(To lze, neboť $p > 0$ a $\varepsilon, \gamma > 0$ jsou libovolně malé. Tím pádem levá strana je libovolně velká) pak dostáváme:

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{2.16}$$

2. $z_i = 1$ a $1 - \gamma < x_i < 1$: φ_i splňuje na $[x_i, 1]$ Lagrangeovu větu o střední hodnotě a tedy platí:

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = p[-(\tilde{x}_i + \varepsilon)^{-p-1} + (1 - \tilde{x}_i + \varepsilon)^{-p-1}]|x_i - z_i|,$$

kde $\tilde{x}_i \in (x_i, 1)$. Zvolme $\gamma < 1/2$ a neboť $p > 0$ pak platí:

$$\begin{aligned}\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) &\geq p[-(1 - \gamma + \varepsilon)^{-p-1} + (1 - (1 - \gamma) + \varepsilon)^{-p-1}]|x_i - z_i| \\ &= p[-(1 - \gamma + \varepsilon)^{-p-1} + (\gamma + \varepsilon)^{-p-1}]|x_i - z_i|.\end{aligned}$$

Pokud použijeme stejné argumenty jako výše včetně 2.15, tak dostaneme nerovnost 2.16.

3. $z_i = x_i = 0$ nebo $z_i = x_i = 1$: dostáváme:

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = 0.$$

Pokud ε, γ splňují 2.15, pak dostáváme:

$$\varphi(\mathbf{x}, \varepsilon) - \varphi(\mathbf{z}, \varepsilon) \geq \tilde{L} \sum_{i=1}^{n} |x_i - z_i| = \tilde{L}\|\mathbf{x} - \mathbf{z}\|_1 \geq \tilde{L}\|\mathbf{x} - \mathbf{z}\|_\infty,$$

pro všechna $\mathbf{z} \in T \cap \{0,1\}^n$ a pro všechny $\mathbf{x}$ takové, že $\|\mathbf{x} - \mathbf{z}\|_\infty < \gamma$.
Definujme $S(\mathbf{z}) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{z}\|_\infty < \gamma\}$ a $S = \cup_{i=1}^{N} S(\mathbf{z}_i)$, kde N je počet bodů $\mathbf{z} \in T \cap \{0,1\}^n$, a tedy dostáváme, že platí 2.12 ve Větě 4.
Nechť pro $\overline{\mathbf{x}} \in \mathbb{R}^n$ platí: $\overline{x_j} = \gamma$ pro právě jedno j a $\overline{x_i} \in \{0,1\}$ pro všechny

$i \neq j$, tím pádem $\overline{\mathbf{x}} \notin S$, dále definujme nekonečnou posloupnost $\{\varepsilon^k\}$ takovou, že $\varepsilon^k \to 0^+$ pro $k \to \infty$, pak pro všechna $\mathbf{z} \in \{0,1\}^n$ platí(p > 0):

$$\lim_{k\to\infty}[\varphi(\overline{\mathbf{x}},\varepsilon^k) - \varphi(\mathbf{z},\varepsilon^k)] =$$
$$\lim_{k\to\infty}[-(\gamma+\varepsilon^k)^{-p} - (1-\gamma+\varepsilon^k)^{-p} + (\varepsilon^k)^{-p} + (1+\varepsilon^k)^{-p}] = \infty,$$

takže předpoklad 2.13 z Věty 4 platí.
Pro všechny $\mathbf{x} \in X \backslash S$ a pro všechny $\varepsilon > 0$ máme:

$$\begin{aligned}&\varphi(\mathbf{x},\varepsilon) - \varphi(\overline{\mathbf{x}},\varepsilon) = \\ &\sum_{i\neq j}[-(x_i+\varepsilon)^{-p} - (1-x_i+\varepsilon)^{-p}] - (n-1)(-(\varepsilon)^{-p} - (1+\varepsilon)^{-p}) \\ &- (x_j+\varepsilon)^{-p} - (1-x_j+\varepsilon)^{-p} + (\gamma+\varepsilon)^{-p} + (1-\gamma+\varepsilon)^{-p} \geq 0,\end{aligned}$$

kde $\gamma \leq x_j \leq 1-\gamma$, neboť platí:

$$\begin{aligned}(x_j+\varepsilon)^{-p} &\leq (\gamma+\varepsilon)^{-p} \\ (1-x_j+\varepsilon)^{-p} &\leq (1-\gamma+\varepsilon)^{-p},\end{aligned}$$

pak tedy platí předpoklad 2.14 z Věty 4.
Tvrzení Věty 5 pro penalizační funkci 2.4, tím pádem plyne z Věty 4.

□

3. Numerická studie

V této kapitole se budeme zabývat minimalizaci rizika při zadaném minimálním očekávaném výnosu portfolia s omezeným počtem aktiv(ang. *Sparse portfolio*). Budeme vycházet z článku [4] a ze skript [5][Kapitola 6.3]. Tento speciální typ portfolia redukuje objem vstupních poplatků aktiv a díky malému rozsahu se dobře spravuje. Tento problém můžeme formulovat například takto:

$$\begin{aligned} \text{minimalizovat} \quad & R(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in X_r \\ & \|\mathbf{x}\|_0 \leq \kappa, \end{aligned} \tag{3.1}$$

kde $\|\mathbf{x}\|_0$ = počet nenulových složek vektoru $\mathbf{x} = (x_1, \ldots, x_n)$. Tedy celkový počet různých aktiv je omezen $\kappa \in \mathbb{N}$ a $\kappa \leq n$. Dále $R(\mathbf{x}) = \mathbf{x}^T\mathbf{V}\mathbf{x}$, kde $\mathbf{V}$ je varianční matice n aktiv. Předpokládáme, že váhy aktiv jsou omezeny minimální očekávanou výnosností portfolia r ve smyslu:

$$X_r = \left\{ \mathbf{x} \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1, \mathbb{E}(\sum_{i=1}^n x_i R_i) \geq r \right\},$$

kde R_i je náhodný výnos aktiva i s konečným druhým momentem. Podmínka $\|\mathbf{x}\|_0 \leq \kappa$ se dá ekvivalentně převést na celočíselné(binární) omezení takto:

$$\begin{aligned} & \sum_{i=1}^n y_i \geq n - \kappa \\ & x_i y_i = 0 \qquad i = 1, \ldots, n, \\ & y_i \in \{0, 1\} \qquad i = 1, \ldots, n. \end{aligned}$$

Binární proměnná y_i je rovna 1 právě tehdy, když příslušná váha $x_i = 0$. Povšimneme si, že podmínka $\sum_{i=1}^n y_i \geq n - \kappa$ zajišťuje, že je nejméně $n - \kappa$ vah, jež jsou nulové(ekvivalentně je nejvýše κ nenulových vah).

Scholtesova regularizace
Pro parametr $t \geq 0$ formulujeme, viz [4][Chapter 2.2.1]:

$$\begin{aligned} \text{minimalizovat} \quad & R(\mathbf{x}) \\ \text{za podmínek} \quad & \mathbf{x} \in X_r \\ & \sum_{i=1}^n z_i \geq n - \kappa \\ & -t \leq x_i z_i \leq t \qquad i = 1, \ldots, n, \\ & 0 \leq z_i \leq 1 \qquad i = 1, \ldots, n. \end{aligned} \tag{3.2}$$

Algoritmus spočívá v postupném řešení úloh se startovacím bodem, který je optimálním řešení v předešlé iteraci, a v získání sekvence bodů, která za předpokladů popsaných v [7], konverguje k optimálnímu řešení úlohy 3.1.

Zadefinujme i penalizační problém vycházející z 3.2 s příslušnou penalizační funkcí 2.3 jako:

$$\begin{aligned} \text{minimalizovat} \quad & R(\mathbf{x}) + \sum_{i=1}^{n} \log(z_i + \varepsilon) + \log((1 - z_i) + \varepsilon) \\ \text{za podmínek} \quad & \mathbf{x} \in X_r, \\ & \sum_{i=1}^{n} z_i \geq n - \kappa, \\ & -t \leq x_i z_i \leq t \qquad i = 1, \ldots, n, \\ & 0 \leq z_i \leq 1 \qquad i = 1, \ldots, n, \end{aligned} \tag{3.3}$$

kde $\varepsilon > 0$ a $t \geq 0$.

Tento penalizační problém použijeme pro spočítání deseti různých minimalizací rizika portfolia s omezeným počtem aktiv. Potřebná data získáme ze stránky [6]. Jedná se o 2x5 portfolií: orl200-005-f, -g, -h, -i, -j a orl300-005-f, -g, -h, -i, -j. Každé portfolio z první sady obsahuje 200 aktiv a z druhé sady 300 aktiv. U každého portfolia z uvedených sad známe varianční matici, vektor náhodných výnosů daných aktiv a minimální očekávaný výnos celého portfolia. Položme $\kappa = 10$. Veškeré výpočty byly spočteny v programu *MATLAB R2018a*. Jako počáteční bod použijeme:

$$x_i = 0 = z_i \qquad i = 1, \ldots, n.$$

Položme penalizační parametr $\varepsilon = 10^{-k}$, kde $k = 3, 4, 5, 6$. Parametr t volíme roven nule. Pro všechny hodnoty penalizačního parametru ε, spočítáme minimalizační penalizační problém 3.3. Výpočty jsou uvedeny v tabulkách 3.1 a 3.2. První tabulka je pro portfolia s 200 aktivy a druhá pro portfolia s 300 aktivy.

Pro portfolia o 200 aktivech bylo dosaženo nejlepších výsledků pro hodnotu $\varepsilon = 10^{-6}$ (3 z 5), avšak u úlohy f bylo dosaženo nejlepších výsledků pro $\varepsilon = 10^{-k}, k = 4, 5, 6$. S klesající hodnotou penalizačního parametru, by se měly zlepšovat(klesat) hodnoty pomocné funkce, neboť zde dochází k větší penalizaci odchylek od binarity. Tento trend je názorněji vidět ve druhé tabulce 3.2. Zároveň pozorujeme, že u nejlepších výsledků bylo dosaženo horní hranice pro parametr κ(Princip diverzifikace).

Pro portfolia o 300 aktivech bylo ve všech portfoliích dosaženo nejlepších výsledků pro hodnotu $\varepsilon = 10^{-6}$. Zde je již více patrný trend klesající hodnoty pomocné funkce, pro klesající hodnotu penalizačního parametru ε. Časová náročnost je zde 3-4-krát větší než u portfolií s 200 aktivy. U každého nejlepšího výsledku bylo dosaženo nejvyšší κ pro dané portfolium(9 či 10). Za povšimnutí stojí, že hodnota $\varepsilon = 10^{-4}$ byla ve všech případech početně výrazně časově náročnější než ostatní hodnoty ε.

ε	10^{-6}	10^{-5}	10^{-4}	10^{-3}
f	**79,84**	**79,84**	**79,84**	350,73
	18,5	17,2	16,9	37,3
	10	10	10	2
g	**57,8**	89,9	107,5	89,9
	47,6	45,1	30,7	15,9
	10	6	5	6
h	56,96	62,32	56,24	**56,08**
	36,5	52,8	45,7	20,4
	10	9	10	10
i	87,06	**78,47**	124,74	148,31
	41,1	31,9	46,3	16,8
	9	10	6	5
j	**77,43**	77,70	107,53	108,86
	30,6	32,7	48,8	18,7
	10	10	7	7

Tabulka 3.1: Numerické výsledky pro portfolia typu orl200-005 - optimální hodnota, čas(sekundy), dosažená κ. Nejlepší výsledky jsou zvýrazněny tučně.

ε	10^{-6}	10^{-5}	10^{-4}	10^{-3}
f	**112,90**	186,27	139,76	139,42
	127,7	116,15	134,5	58,3
	10	6	8	8
g	**125,03**	127,72	158,41	139,56
	130,4	116,9	130,8	58,5
	9	9	7	8
h	**113,71**	126,63	140,30	219,24
	120,5	121,8	135,6	53,2
	10	9	8	5
i	**111,62**	126,25	182,15	216,45
	126,5	120,5	136,2	53,1
	10	9	6	5
j	**126,18**	140,50	139,00	159,09
	116,6	118,7	175,8	88,3
	9	8	8	7

Tabulka 3.2: Numerické výsledky pro portfolia typu orl300-005 - optimální hodnota, čas(sekundy), dosažená κ. Nejlepší výsledky jsou zvýrazněny tučně.

Závěr

V práci jsme si představili jeden z možných přístupu řešení nelineárních optimalizačních úloh a to jest převod na úlohu hledaní volného extrému, kde jsou omezení převedena do účelové funkce pomocí takzvané penalizační funkce. Seznámili jsme se s Metodou vnějšího bodu, která řeší původní úlohu pomocí postupně zvětšujícího se penalizačního parametru. Byl uveden i algoritmus pro řešení úlohy pomocí Metody vnějšího bodu a jeho demonstrace na příkladu, kde jsme pozorovali chování penalizační funkce a konvergenci postupných řešení k optimální hodnotě s rostoucím penalizačním parametrem. Poté jsme se věnovali penalizačním funkcím, které nevyžadují limitní přiblížení penalizačního parametru k nekonečnu, jsou to takzvané Exaktní penalizační funkce. Následně se práce začala zaobírat celočíselným binárním nelineárním programováním, kde bylo uvedeno několik penalizačních funkcí, jež jsou vhodné pro řešení tohoto binárního problému. Zde byla také uvedena analogie důkazu pro konkrétní penalizační funkci, vycházející ze základních kroků důkazu pro obecné penalizační funkce tohoto typu. V numerické části se práce zabývala minimalizací rizika při zadaném minimálním očekávaném výnosu portfolia s omezeným počtem aktiv. Uvedli jsme si základní problém vycházející z Markowitzova modelu, který byl následně převeden, pomocí předchozích znalosti, na penalizační problém. Poté jsme si uvedli dva typy sad portfolií s omezeným počtem aktiv, které byly řešeny pomocí předchozího penalizačního problému. Studovali jsme dosažené výsledky řešení daného portfolia při změně penalizačního parametru.

Seznam použité literatury

[1] Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. *Nonlinear Programming: Theory and Algorithms.* Wiley-Interscience, Hoboken, N.J., 3rd ed., 2006, 204-205, 469-501.

[2] Lucidi, S., Rinaldi, F. *Journal of Optimization Theory and Applications.* Springer Science+Business Media, 2010, 479-488.

[3] Pick L., Hencl S., Spurný J., Zelený M.(2019) *Matematická analýza 1.* http://www.karlin.mff.cuni.cz/~pick/analyza.pdf

[4] Branda, M. *Sparsity and regularization in portfolio selection problems.* Managing and Modelling of Financial Risks, Ostrava, 2018, 45-52.

[5] Dupačová J., Lachout P. *Úvod do optimalizace.* MatfyzPress, Matematicko-fyzikální fakulta - Univerzita Karlova, 2011, 58-59.

[6] Frangioni, A., Gentile, D.C. *Mean-variance problem with minimum buy-in constraints, data and documentation.* http://www.di.unipi.it/optimize/Data/MV.html

[7] Branda, M., Bucher, M., Červinka, M., Schwartz, A.: *Convergence of a Scholtes-type regularization method for cardinality-constrained optimization problems with an application in sparse robust portfolio optimization.* Computational Optimization and Applications 70 (2), 2018, 503-530.