UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

KALIBRÁCIA MODELOV FINANČNEJ MATEMATIKY POUŽITÍM DVOJKRITERIÁLNEHO GREY WOLF ALGORITMU

DIPLOMOVÁ PRÁCA

2023 Bc. Mária RUDOLFOVÁ

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

KALIBRÁCIA MODELOV FINANČNEJ MATEMATIKY POUŽITÍM DVOJKRITERIÁLNEHO GREY WOLF ALGORITMU

DIPLOMOVÁ PRÁCA

Študijný program:	ekonomicko-finančná matematika a modelovanie
Študijný odbor:	matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. RNDr. Beáta Stehlíková, PhD.

Bratislava 2023 **Bc. Mária RUDOLFOVÁ**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Mária Rudolfová
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Kalibrácia modelov finančnej matematiky použitím dvojkriteriálneho grey wolf algoritmu
Calibration of financial mathematics models using bicriteria grey wolf algorithm

Anotácia: V diplomovej práci Patrície Dianovej (2020) bola v jazyku R naprogramovaná 'grey wolf' metóda pre riešenie dvojkriteriálnej optimalizácie, ktorá bola navrhnutá v literatúre a dostupná v iných jazykoch. Úspešne sa použila na kalibráciu úrokových mier. Cieľom tejto diplomovej práce je nadviazať na tento výsledok - navrhnúť a zrealizovať kalibráciu parametrov ďalších finančných modelov.

Vedúci: doc. RNDr. Beáta Stehlíková, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 11.01.2022

Dátum schválenia: 14.01.2022 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Poďakovanie Týmto sa chcem poďakovať svojej vedúcej diplomovej práce doc. RNDr. Beáte Stehlíkovej PhD., za jej nekonečnú trpezlivosť, cenné rady a odborné pripomienky, ktoré mi pomohli dokonvergovať túto prácu k odovzdaniu. Veľká vďaka patrí celej mojej rodine, ktorá nikdy o mne nezapochybovala a vždy verila, že všetko s Božou pomocou nejako zvládnem.

Abstrakt v štátnom jazyku

RUDOLFOVÁ, Mária: Kalibrácia modelov finančnej matematiky použitím dvojkriteriálneho grey wolf algoritmu [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. RNDr. Beáta Stehlíková, PhD., Bratislava, 2023, 81s.

Hlavným cieľom práce bude nájsť čo najlepšie parametre pre fitovanie poistných a finančných dát využitím genetického dvojkriteriálneho grey wolf algoritmu. V úvode priblížíme myšlienku algoritmu, na základne ktorej hľadáme optima funkcií. Zároveň vysvetlíme pojem Pareto optimálnych riešení. V aplikačnej časti budeme chcieť nájsť čo najlepšie rozdelenie, ktoré by popisovalo dáta predstavujúce škodové udalosti vzniknuté pri požiaroch v Dánsku. V ďalšej aplikácii porovnáme fitovanie výnosov trhového indexu $S\&P$ 500 Normálnym a Variance gamma rozdelením. Taktiež nájdeme optimálne váhy portfólia štyroch akcií pre Markowitzov a Mean-Gini model. V poslednej časti práce nakalibrujeme Vašíčkov model pre Európske úrokové miery. Úrokové miery Českej republiky nakalibrujeme s predpokladom, že sa budú priťahovať k Európskym.

Kľúčové slová: Grey wolf algoritmus, Burrovo rozdelenie, LogPH rozdelenie, Normálne rozdelenie, Variance gama rozdelenie, Markowitzov model, Mean-Gini model, úrokové miery

Abstract

RUDOLFOVÁ, Mária: Calibration of financial mathematics models using bicriteria grey wolf algorithm [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. RNDr. Beáta Stehlíková, PhD., Bratislava, 2023, 81p.

The main objective of the work will be to find the best possible parameters for fitting insurance and financial data by using the genetic two-criteria grey wolf algorithm. In the introduction, we will introduce the idea of the algorithm, based on which it searches for the optima of the functions. We also explain the concept of Pareto optimal solutions. In the application part, we will want to find the best possible distribution to describe the data representing damage events arising from fires in Denmark. In the next application, we will compare the fitting of the returns of the market index $S\&P$ 500 by the Normal and Variance Gamma distributions. We also find the optimal portfolio weights of the four stocks for the Markowitz and Mean-Gini models. In the last part of the paper, we calibrate the Vasicek model for European interest rates. We calibrate the interest rates of the Czech Republic with the assumption that they will converge to the European ones.

Keywords: Grey wolf algorithm, Burr distribution, LogPH distribution, Normal distribution, Variance gamma distribution, Markowitz model, Mean-Gini model, interest rates

Obsah

Úvod

Financie, bankovníctvo, investovanie či poisťovníctvo. Tieto a mnoho ďalších veľkých a komplexných oborov vyžadujú prácu matematikov na dennej báze. Pracujú s modelmi, ktoré by boli schopné čo najlepšie popísať aktuálnu situáciu a zároveň pripraviť čo najlepšiu predikciu do budúcnosti. Jedným z dobrých nástrojov, ako to dosiahnuť sa v realite javia byť parametrické modely.

Parametrický model je vcelku veľmi jednoduchý nástroj pre modelovanie. Už z názvu je zrejmé, že pre modelovanie dát budeme používať model, ktorý je definovaný parametrami. Napríklad hustota normálneho rozdelenia je definovaná parametrami μ a σ, kedy po zadaní správnych parametrov dostávame hustotu, ktorá popisuje normálne dáta. Najťažšou otázkou však ostáva, ako dané parametre v modeli zvoliť?

Na túto otázku sa budeme snažiť zodpovedať v tejto diplomovej práci. Patrícia Dianová vo svojej diplomovej práci [14] naprogramovala genetický dvojkriteriálny algoritmus pre hľadanie Pareto optimálnych riešení.

V úvode pripomenieme čo so sebou prináša multikriteriálna optimalizácia. Následne priblížime stručnú myšlienku algoritmu, čo je jeho nosnou ideou pri hľadaní optima. Pre lepšie pochopenie a popis algoritmu nám pomôže [19] [26] [27].

V prvej časti práce využitím daného algoritmu odhadneme parametre pre dáta, ktoré popisujú výšky vzniknutých škôd pri požiaroch v Dánsku. Najprv nakalibrujeme Burrovo rozdelenie, ktoré sa veľmi často používa v poisťovníctve pri modelovaní škôd. Vychádzajúc z [1] a [2] odhadujeme parametre pre $logPH$ rozdelenie. V [2] nachádzame optimálne parametre pre $logPH$ rozdelenie pozorovaných dát, získané inou optimalizáciou, akú používame my. To vnímame ako veľkú výhodu, kvôli kontrole a porovnaniu nami dosiahnutých optimálnych parametrov rozdelenia. Keďže algoritmom nachádzame Pareto optimálne riešenia, požadujeme, aby aspoň jedno z optimalizačných kritérií bolo pre získané riešenia menšie alebo rovné ako pre riešenie z článku.

Asi najznámejší model pre oceňovanie akcií (Black - Scholesov model) pracuje so silným predpokladom o normalite dát. My sa bližšie pozrieme na rozdelenie výnosov trhového indexu $S\&P$ 500. Ukážeme, že dáta majú od normálneho rozdelenia naozaj ďaleko. Posnažíme sa modelom čo najlepšie odhadnuť parametre pre normálne rozdelenie a vychádzajúc z [9] ukážeme, že Variance gama rozdelenie, spomedzi nami vybraných rozdelení, naozaj

najlepšie popisuje dáta. Oprieme sa o [4] a [9].

Markowitzov model je jeden z najfrekventovanejších prístupov pre nájdenie optimálnych váh pri zostavovaní portfólia. Využíva veľmi jednoduchú a logickú myšlienku, kedy chceme minimalizovať riziko portfólia prezentované varianciou portfólia a následne maximalizovať priemerný ročný výnos akcií. Využitím algoritmu sa pokúsime nájsť optimálne váhy pre zostavenie portfólia z najvýnosnejších akcií trhového indexu $S\&P$ 500, a to akcie Applu, Microsoftu, Amazonu a Invidii. Riziko portfólia prezentované varianciou hodnoty portfólia vymeníme za Mean-Gini kritérium a pozrieme sa, čo sa zmení pri takejto zmene optimalizácie.

V poslednej časti práce nakalibrujeme dlhodobé európske úrokové miery a úrokové miery Českej republiky. Využitím Vašíčkovho jednofaktorového modelu odhadneme parametre pre európske úrokové miery. Predpokladáme, že úrokové miery Českej republiky budú priťahované k európskym, keďže aj Česká republika je členom Európskej únie a menová politika krajiny je ovplyvňovaná Európskou centrálnou bankou.

V každej aplikačnej časti podrobne vysvetlíme a popíšeme, aké dve funkcie (kritéria) budú vstupovať do algoritmu ako optimalizačné funkcie.

1 Úvod do problematiky

Na úvod sa pozrieme na príklady, kedy sa využívala multikriterálna optimalizácia. Keďže jedna úloha nám (pre multikriteriálnu optimalizáciu) môže priniesť viacej riešení, ozrejmíme čitateľovi ako správne chápať takto získané Pareto optimálne výsledky. Na záver v skrátke popíšeme algoritmus, s ktorým budeme pracovať. Taktiež prinášame prehľad, kde bol náš grey wolf algoritmus použitý.

1.1 Multikriteriálna optimalizácia

Oblasť financií, ako aj veľmi blízka oblasť poisťovníctva predstavujú odvetvia, kde spraviť rozhodnutie týkajúce sa financií, rozpočtov, či stanovení rezerv nie je priamočiara a jednoduchá záležitosť. Finančníci pri modelovaní potrebujú odzrkadliť hlbokú neistotu, dynamické prostredie trhov, finančné problémy a v neposlednom rade subjekty, ktoré sa podieľajú na vývoji modelu (manažér, regulátor a pod.). Tu zohrávajú dôležitú úlohu optimalizačné modely s viacerými kritériami, keďže jedno kritérium nedokáže dostatočne popísať komplexnosť problematiky. Pozrieme sa na rôzne príklady, kde sa miltikriterialna optimalizácia využíva.

Ako príklad si môžeme vziať výber portfólia. Pri takomto probléme by sme chceli uspokojiť dva hlavné ciele. Týmito cieľmi sú výnos a riziko. Takémuto dvojkriteriálnemu optimalizačnému problému sa venujú v mnohých publikáciach. V [24] na riešenie optimalizačného problému predstavujú evolučný algoritmus DEMPO (Differential Evolution for Multiobjective Portfolio Optimalization). Tento algoritmus následne porovnávajú s kvadratickou optimalizáciou a ďalším evolučným algoritmom určeným na multikriteriálnu optimalizáciu NSGA-II.

V ďalšej publikácii [15] sa autori zamerali na multikriteriálnu optimalizáciu v odvetví financií a investovania. Nachádzajú sa tu optimalizačné problémy týkajúce sa aj optimalizácie portfólia. Model konštruuje portfóliá tak, aby boli čo najlepšie diverzifikované v rámci vybraných aktív a zároveň boli globálne optimálne s vhodnou mierou výnosnosti. Taktiež sa tu venujú aj trojkriteriálnej optimalizácii, kde cieľom kritérií je zachytiť očakávaný výnos, averziu voči strate a mieru rizika.

Vo viackriterialnej optimalizácii je potrebné zadefinovať a ujasniť si, aké kritéria bu-

deme používať pri kalibrovaní modelov. Jedno z najčastejšie používaných kritérií je neodmysliteľne maximalizácia funkcie vierohodností (maximum likelihool function). Ďalším populárnym optimalizačným kritériom je nájdenie minima chybovej funkcie (minimum error function), resp. minima kvadratických odchýlok medzi fitovanými dátami ($\hat{x}$) vybraným rozdelením s parametrom θ a pozorovanými dátami (x), t.j. $\min ||x - \hat{x}||^2$. Použitie týchto kritérií pri mltikriteriálnej optimalizácií nachádzame napríklad aj v [31].

1.2 Pareto optimálne riešenia

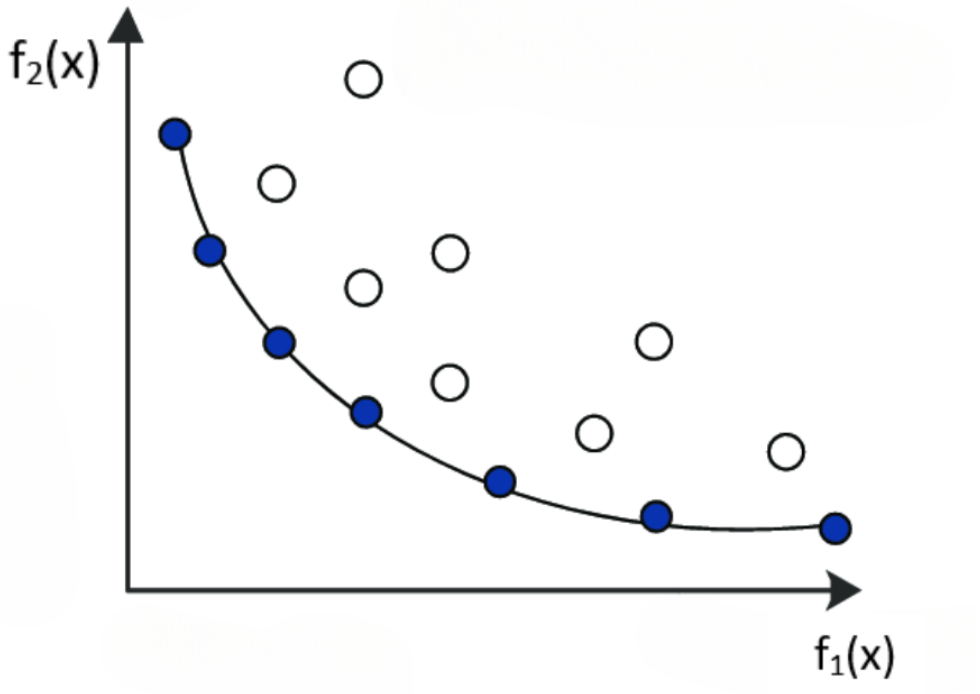
Pareto optimálne výsledky dostávame pri multikriteriálnej optimalizácii. Teoretické spracovanie a zároveň praktické aplikácie nachádzame v [23] alebo v [28], kde sa Pareto optimálne riešenia aplikujú v chemickom inžinierstve.

Matematicky by sme problematiku multikriterálnej optimalizácie vedeli zapísať nasledovne:

$$\begin{aligned} \max \quad & F(\vec{x}) = f_1(\vec{x}), f_2(\vec{x}), \ldots, f_o(\vec{x}) \\ & g_i(\vec{x}) \geq 0,\ i = 1, \ldots, m \\ & h_i(\vec{x}) = 0,\ i = 1, \ldots, p \\ & L_i \leq x_i \leq U_i,\ i = 1, \ldots, n, \end{aligned}$$

kde n je počet premenných, o je počet funkcií, ktoré chceme zoptimalizovať, m je počet ohraničení na nerovnosti prezentované funkciou $g_i(x)$ a p je počet ohraničení na rovností znázornené funkciou $h_i(x)$. Premenná x_i patrí do intervalu $[L_i, U_i]$ [27].

Je dôležité poznamenať, že pri multikriterálnej optimalizácii nezískavame lokálne (resp. globálne) optimum. Získavame Pareto optimálne riešenia (nedominované riešenia). To znamená, že pre dané riešenie nedokážeme znížiť hodnotu jednej účelovej funkcie bez toho, aby sme nezvýšili hodnotu druhej účelovej funkcie. Takto získané riešenia tvoria množinu Pareto optimálnych riešení. Obr.1 znázorňuje ako môžu vyzerať Pareto optimálne riešenia - nedominované riešenia (modrou farbou) pri optimalizácii dvoch funkcií $f_1(x)$ a $f_2(x)$. Takéto riešenia ležia na krivke, ktorá sa nazýva Pareto optimálny front. Bielé krúžky znazorňujú dominované riešenia optimalizačného problému. [14]



Obr. 1: Dominované, nedominované riešenia a Pareto front pri optimalizácii dvoch funkcií $f_1(x)$ a $f_2(x)$ [10]

1.3 *Grey wolf* algoritmus

Detailný popis používaného algoritmu načrtla vo svojej práci [14] Patrícia Dianová. Využívame článok [26], v ktorom je detailne popísaný genetický Grey wolf algoritmus pre optimalizáciu úlohy s jednou optimalizačnou funkciou. V [27] je algoritmus upravený pre dvojkriteriálnu optimalizáciu. Viac informácii o grey wolf optimalizácii, ale aj o iných genetických algoritmoch nájdeme aj online [19].

My sa posnažíme len v skratke ozrejmiť hlavnú myšlienku daného algoritmu. Pokladáme to za dôležité, keďže na danom algoritme je postavená celá práca.

Grey wolf algoritmus je genetický algoritmus navrhnutý tak, aby dokázal nájsť optimálne riešenia daného systému. Myšlienka hľadania optima je na základe pozorovania správania sa svorky sivých vlkov voľne v prírode. V nasledujúcich častiach si v skratke popíšeme, ako sa táto myšlienka matematicky sformulovala a následne použila v optimalizácii.

Šedé vlky sa vo voľnej prírode riadia striktnou hierarchiou, kde nachádzame vodcov svorky (tzv. alfa vlkov), ich zástupcov (bety) a ostatné vlky. Tak aj pre nás, najlepšieho

kandidáta na riešenie bude označovať gréckym symbolom α, druhé a tretie najlepšie riešenia symbolmi β (resp. δ). Grécke písmeno ω bude označovať zvyšné kandidátske riešenia optimalizačného problému. Systém predpokladá, že α, β a δ predstavujú najlepšie riešenia, a preto podľa nich aktualizujú svoju pozíciu zvyšné kandidátske riešenia (ω).

Nájdenie optima sa odohráva v troch bodoch, ktoré sú inšpirované útokom šedých vlkov na korisť.

1. Obkľúčenie koristi

   Pod pojmom korisť budeme rozumieť nami hľadané optimum a vlk predstavuje kandidátske riešenie. V tejto časti sa prepočítava pozícia vlka podľa toho, kde sa nachádza korisť (hľadané optimum). Nachádzajú sa tu parametre, ktoré zaznamenávajú náhodnosť (r_1, r_2), parameter a (lineárne sa zmenšuje od 2 k 0), ktorý reflektuje približovanie sa ku koristi, vektor $\vec{X}$ hovorí o pozícii vlka (kandidátskeho riešenia), $\vec{X_p}$ predstavuje pozíciu koristi (hľadaného optima). Zmenu pozície vlka v čase počítame nasledovne:

$$\vec{D} = |\vec{C}\vec{X_p}(t) - \vec{X(t)}|, \tag{1}$$

$$\vec{X(t+1)} = \vec{X_p(t)} - \vec{A}\vec{D}, \tag{2}$$

$$\vec{A} = 2\vec{a}\vec{r_1} - \vec{a}, \tag{3}$$

$$\vec{C} = 2\vec{r_2}. \tag{4}$$

2. Lov

   Ako sme spomínali, α, β a δ sú nosičmi pre nás zatiaľ najlepších dosiahnutých riešení. Preto tieto pozície si uložíme a ostatné vlky (zvyšné kandidátske riešenia) ω sa prepočítajú vzhľadom na doteraz najlepšie získané riešenia.

$$\vec{D_\alpha} = |\vec{C_1}\vec{X_\alpha} - \vec{X}|, \tag{5}$$

$$\vec{D_\beta} = |\vec{C_2}\vec{X_\beta} - \vec{X}|, \tag{6}$$

$$\vec{D_\delta} = |\vec{C_3}\vec{X_\delta} - \vec{X}|, \tag{7}$$

   kde C_1, C_2 a C_3 sa napočítavajú vyššie spomenutým postupom (4). X_α, X_β a X_δ sú tri doposiaľ najlepšie získané riešenia v iterácii t. Následne sa aktualizuje pozícia

kandidátskych riešení pre čas $t+1$

$$\vec{X_1} = \vec{X_\alpha} - \vec{A_1}\vec{D_\alpha}, \tag{8}$$

$$\vec{X_2} = \vec{X_\beta} - \vec{A_2}\vec{D_\beta}, \tag{9}$$

$$\vec{X_3} = \vec{X_\delta} - \vec{A_3}\vec{D_\delta}, \tag{10}$$

$$\vec{X(t+1)} = \frac{\vec{X_1} + \vec{X_2} + \vec{X_3}}{3}. \tag{11}$$

Parameter A_1, A_2 a A_3 sa prepočítava podľa (3).

3. Útok

   Náhodný parameter A hovorí, či sa s kandidátmi blížime ku skutočnému riešeniu (ku koristi), alebo sa od neho oddiaľujeme. $\vec{A} \in [-2a, 2a]$ kde $a \in [0, 2]$. Ak $|\vec{A}| < 1$, vtedy kandidátske riešenia smerujú k optimu. Naopak, ak by $|\vec{A}| > 1$, tak sa snažíme kandidátske riešenia odkloniť a nájsť niečo lepšie.

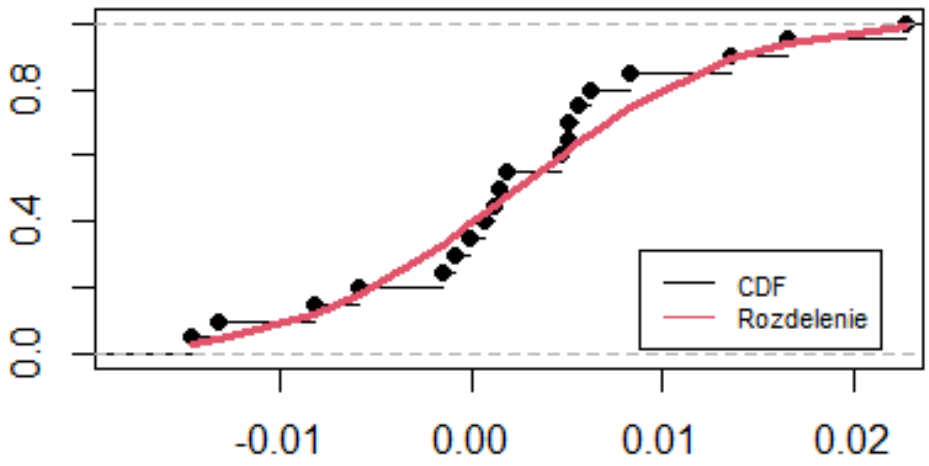
V našej práci budeme využívať dvojkriteriálny grey wolf optimalizačný algoritmus, kde sa budeme snažiť nájsť Pareto optimálne riešenia pre dve optimalizačné funkcie. Hlavná myšlienka hľadania optima ostáva nemenná (ako sme popísali vyššie). Pochopiteľne sa tu budú nachádzať aj odlišnosti. Budeme potrebovať archív, kvôli ukladaniu doteraz získaných nedominovaných riešení. Tento archív sa bude riadiť podmienkami, ktoré budú hovoriť, či už získané riešenia budú prepísané novými nedominovanými riešeniami, či sa získané riešenie uloží do archívu ako ďalšie nové nedominované riešenie, alebo sa neudeje nič. Ďalšou novinkou (vzhľadom na jednokriteriálnu optimalizáciu) je stratégia výberu lídra. Má domôcť k optimálnemu výberu najlepších pozícií výsledkov α, β a γ. Snaha je vyberať riešenia v čo najmenej preplnených segmentoch vyhľadávania. Vyhľadaváci priestor je rozsegmentovaný na hyperkocky. Výber sa koná metódou rulety s pravdepodobnosťou pre každú hyperkocku $P_i = \frac{c}{N_i}$, kde $c > 1$ je konštanta a N_i je počet Pareto optimálnych riešení v i-tom segmente. Pravdepodobnosť výberu je vyššia pre menej preplnené segmenty. Bližšie sa tomuto venovali v práci [14].

Daný algoritmus sa využil pre riešenie optimalizačných problémov na množstvo aplikácií, ako napríklad v sieťovom inžinierstve [3] [35], inžinierstve [5] [8], výbere webovej služby [7], energetike [16] [20] a mnoho ďalších. V [14] aplikovali algoritmus na kalibráciu finančného modelu pre modelovanie úrokových mier. My sa posnažíme daným algoritmom

čo najlepšie nafitovať dáta pre rôzne modely v oblasti poisťovníctva (modelovanie rozdelenia dát popisujúcich výšku škodových udalostí) a finančníckeho sveta (rozdelenie výnosu akcií, finančné modely či optimalizácia portfólia).

2 Parametre pravdepodobnostných rozdelení

V nasledujúcej časti práce sa posnažíme nájsť čo najlepší odhad parametrov rôznych tipov rozdelení tak, aby čo najlepšie popisovali dáta, pre ktoré boli vybrané. K takémuto cieľu si najprv potrebujeme zadefinovať kritériá, na základe ktorých budeme optimalizačným algoritmom hľadať parametre pravdepodobnostných rozdelení. Ako sme už skôr spomenuli, za jedno z najčastejších kritérií sa volí vierohodnostná funkcia (likelihood function) alebo chybová funkcia (error function). Inšpirovali sme sa modelovaním v poisťovníctve [29] a ako kritéria pri optimalizácii sme zvolili testové štatistiky.



Obr. 2: Porovnanie empirickej kumulatívnej distribučnej funkcie s distribúciou rozdelenia

Na Obr.2 vidíme čiernou farbou schodovitú funkciu, ktorá predstavuje empirickú kumilatívnu distribučnú funkciu. Našou kalibráciou by sme sa chceli zvoliť parametre distribučnej funkcie rozdelenia (červená krivka) tak, aby sme sa čo najviac priblížiť k danej kumulatívnej distribučnej funkcii. Jednou z mnohých možností, ako zabezpečiť čo najlepšej zhody daných kriviek, by sa dalo zapísať ako optimalizačný problém, kedy by sme chceli minimalizovať najväčšiu odchýlku daných funkcií v dátových bodoch. Presne takúto myšlienku zachytáva Kolmogorovova-Smirnovova testová štatistika (K-S testová štatistika).

Ako prvé kritérium sme zvolili Kolmogorovovu-Smirnovovu testovú štatistiku a druhé kritérium prislúchalo Cramér-von Misesovej testovej štatistike. Matematické vyjadrenie Kolmogorovovej-Smirnovovej testovej štatistiky je nasledovné [17]

$$D = \sup_{x \in \mathcal{D}(F)} |F_n(x) - F(x;\theta)|, \tag{12}$$

kde *sup* v danom predpise označuje suprémum, $F_n(x)$ predstavuje kumulatívnu distribučnú funkciu v bode x a $F(x;\theta)$ označuje distribučnú funkciu rozdelenia v bode x, pre ktoré sa snažíme odhadnúť parametre. Parametre daného rozdelenia pre dáta x sú všeobecne označené gréckym písmenom θ.

Cramér-von Misesova testová štatistika (C-vM testová štatistika) taktiež porovnáva rozdiel medzi pozorovanou a kumulatívnou distribučnou funkciou. Pre pozorované hodnoty x_i vieme danú testovú štatistiku vyjadriť ako [17]

$$T = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F(x_i;\theta) \right], \tag{13}$$

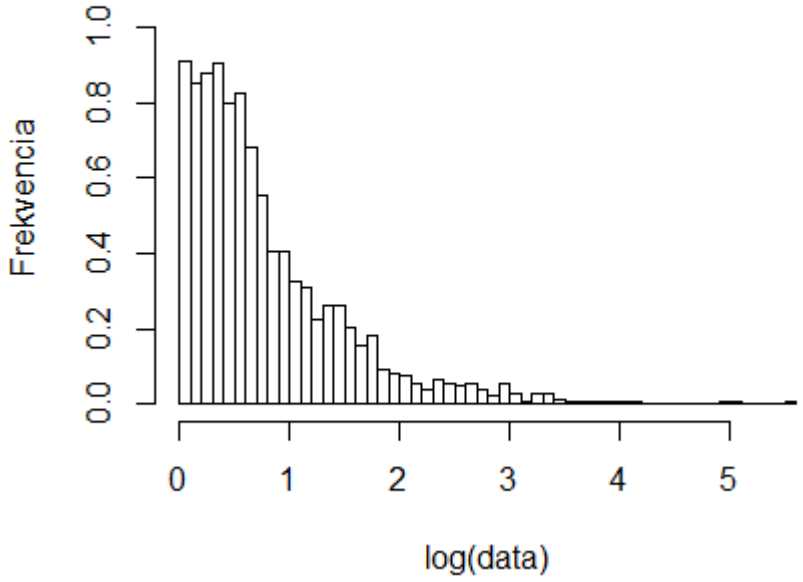
kde opäť $F(x_i;\theta)$ predstavuje distribučnú funkciu daného rozdelenia v i-tom bode s vektorom parametrov θ a n je počet pozorovaní.

2.1 Škodové udalosti pri požiaroch v Dánsku

Z historických dát poistných udalostí sa dá vidieť, že dáta sa vyznačujú asymetriou, ťažkými chvostami. Je zrejmé, že modelujeme nezáporné dáta, keďže modelujeme výšky škôd. Často je dôležité vybrať rozdelenie tak, aby aj chvost rozdelenia čo najlepšie kopíroval dáta. Je pravda, že straty na chvoste sa dejú zriedkavo (s nižšou pravdepodobnosťou). Faktom však ostáva, že keby sme s týmito stratami nepočítali, môžu doviesť poisťovňu do veľkých finančných ťažkostí. Za jedny z často používaných parametrických rozdelení v poisťovníctve pre modelovanie strát môžeme považovať rozdelenia s ťažkými chvostami ako napríklad Weibullovo rozdelenie, Loggama, či Burrovo rozdelenie. V [2] autori úspešne použili pre kalibráciu škodových udalostí LogPH rozdelenie.

Vychádzajúc z [2] sa budeme snažiť v tejto časti čo najlepšie odhadnúť parametre $LogPH$ rozdelenia tak, aby čo najlepšie popisovalo správanie sa našich dát. Autori podrobne popísali dané rozdelenie, vyvodili momenty a popísali jeho správanie. Na praktickom príklade odhadli parametre pre dané rozdelenie využitím $EMPHT$ algoritmu. My sa parametre posnažíme odhadnúť využitím grey wolf algoritmu.

Zozbierané dáta [12] popisujú škody spôsobené požiarom v Dánsku medzi rokmi 1980 až 1990 vrátane hraničných rokov. Spolu máme 2167 pozorovaní. Dáta sú upravené o infláciu a vyjadrené v miliónoch dánskych korún. Pre lepšiu predstavu prikladáme histogram daných dát na Obr.3. Na dáta sme použili logaritmickú funkciu, ktorá nám ich preškáluje kvôli lepšej vizualizácii. Dá sa vidieť, že najnižšia pozorovaná hodnota dát je 1. To je veľmi dôležité kvôli definovaniu distribučnej funkcii $LogPH$ rozdelenia.



Obr. 3: Histogram logaritmu dát popisujúcich výšky škôd pri požiaroch v Dánsku

Zároveň prikladáme základné štatistiky o pozorovanom dátovom súbore. Z týchto štatistík a z histogramu sa dá vidieť, že rozdelenie, ktoré by dobre kopírovalo dáta nebude symetrické, a chceme, aby bolo definované pre kladné dáta.

Minimum	**1. Kvantil**	**Median**	**Priemer**	**3. Kvantil**	**Maximum**
1	1.321	1.778	3.385	2.967	263.250

Tabuľka 1: Základné štatistiky pre dáta popisujúce výšky škôd pri požiaroch v Dánsku

2.1.1 Burrovo rozdelenie

Ako prvé nakalibrujeme Burrovo rozdelenie. Je to rozdelenie, ktoré sa často používa na modelovanie škodových udalostí. Je definované pre kladné hodnoty s tromi parametrami. Hustotu rozdelenia definujeme ako

$$f(x;\alpha,\gamma,\theta) = \frac{\alpha\gamma(x/\theta)^{\gamma}}{x[1+(x/\theta)^{\gamma}]^{\alpha+1}}, \tag{14}$$

kde α, γ, θ sú kladné parametre.

V optimalizačnom algoritme sme nastavili veľkosť populácie vlkov $greyWolves\_num = 500$, a keďže predpokladáme nezápornosť parametrov, ohraničili sme ich intervalom $[0.01, 5]$. Kebyže sa výsledné parametre vyskytovali v blízkosti krajných bodov intervalu ohraničenia, interval by sme rozšírili a porovnali získané riešenia.

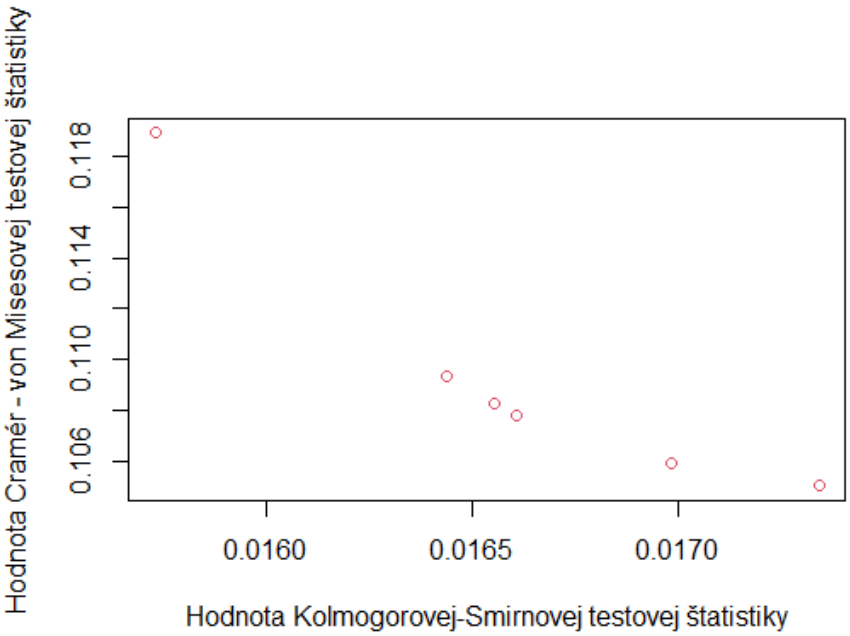
Ďalšie parametre algoritmu potrebné pre výpočet sú nastavené následovne: $MaxIt = 10$ predstavuje maximálny počet iterácií, $Archive\_size = 500$ (archív kvôli ukladaniu Pareto optimálnych riešení), $\alpha_{GW} = 0.1$, $\beta_{GW} = 4$, $\gamma_{GW} = 2$ čo predstavuje štartovacie pozície „vlkov“ (kandidátskych riešení).

Algoritmom sme získali 6 Pareto optimálnych riešení pre dané rozdelenie na dánskych dátach. Na Obr.4 sú znazornené hodnoty optimalizačných funkcií pre dané riešenia, kde x-ová os predstavuje hodnoty K-S testovej štatistiky a y-ová os predstavuje C-vM testovú štatistiku. Hodnoty Kolmogorovovej - Smirnovovej testovej štatistiky sa pohybujú v rozmedzí od 0.01573 po 0.01734. Cramér - von Misesová testová štatistika nadobúda hodnoty od 0.1050 po 0.1189.

Na (Tab.2) znázorňujeme získané parametre Burrovho rozdelena pre Pareto optimálne riešenia α, γ a θ.

Zo získaných riešení sme vybrali také, kde Kolmogorovova - Smirnovova testová štatistika nadobúda najnižšiu hodnotu (krivka pre Pareto optimálne riešenie na grafoch červenou farbou) a najvyššiu hodnotu (znázornené modrou farbou). Obr.5 znázorňuje, ako dobre kopíruje nakalibrovaná distribučná funkcia Burrovho rozdelenia distribučnú funkciu (čiernou farbou).

Na Obr.6 pozorujeme, ako hustoty získané nakalibrovanými parametrami popisujú histogram našich dát.



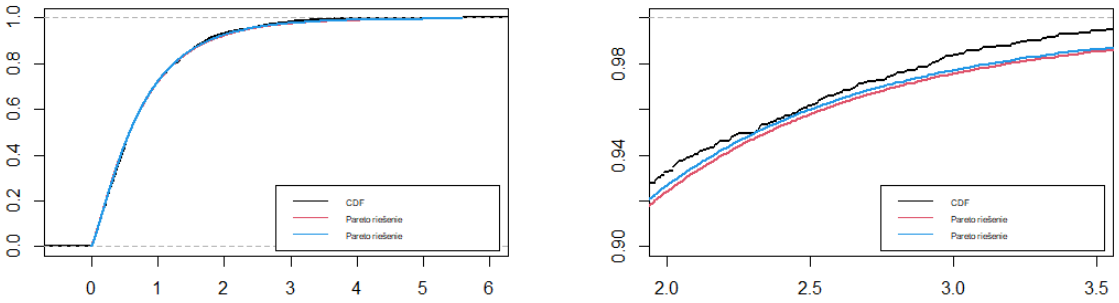
Obr. 4: Získané hodnoty optimalizačných funkcií pre Pareto optimálne riešenia Burrovho rozdelenia

Riešenie	α	γ	θ
1	4.705404	1.231365	2.583958
2	4.918812	1.238368	2.653628
3	4.816350	1.241680	2.604552
4	4.834941	1.239775	2.616517
5	4.797959	1.236336	2.609653
6	4.813522	1.237138	2.614046

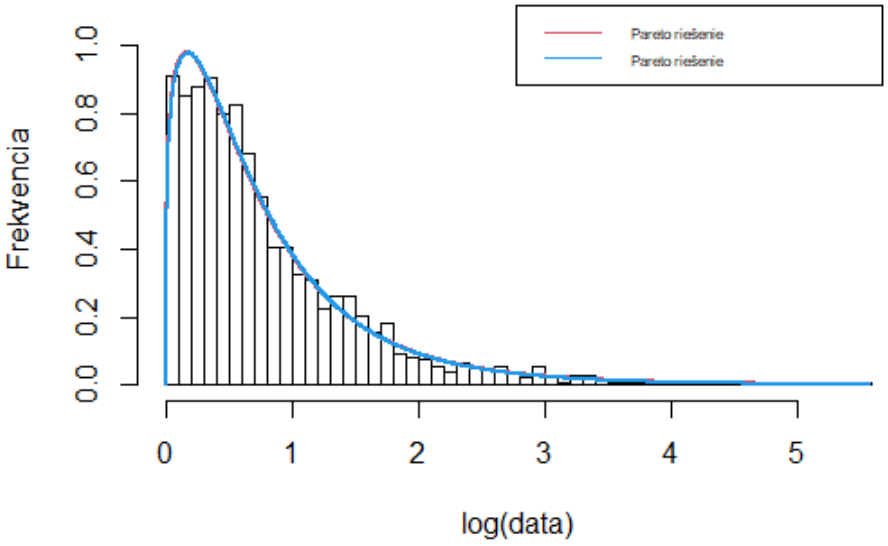
Tabuľka 2: Výsledné parametre Burrovho rozdelenia

2.1.2 LogPH rozdelenie

Parametrické rozdelenie *LogPH* je rozdelením, ktoré bolo úspešne použité pre popis dát, ktoré reflektujú škody spôsobené požiarom v Dánsku (ako sme spomínali vyššie). Dané rozdelenie predstavuje zovšeobecnené Paretovo rozdelenie a sú nám známe vlastnosti hustoty. Taktiež dobre popisuje údaje s ťažkými chvostami (čo je jedným z našich cieľov). Model daného rozdelenia sa snaží čo najlepšie popísať dáta od začiatku až do konca (ne-



Obr. 5: Nakalibrované distribučné funkcie s empirickou kumulatívnou distribučnou funkciou spolu s detailným priblížením



Obr. 6: Hustota Burrovho rozdelenia popisujúca histogram dát

sústredí sa len na chvost, ale na celý interval dát). Ako ďalšiu výhodu vnímame, že pre charakterizáciu hustoty, či distribučnej funkcie máme jeden konkrétny prepis funkcie. Ako nevýhodu vnímame množstvo parametrov, ktoré sú pre model potrebné.

Rozdelenie označujeme $LogPH(\alpha, T)$, kde α a T sú parametrami rozdelenia. Hovoríme, že Y má $LogPH$ rozdelenie, ak $Y = \exp(X)$, kde X má PH rozdelenie (fázové rozdelenie, ang. phase-type distribution).

Distribučná funkcia $LogPH$ rozdelenia je definovaná len pre dáta $x \geq 1$ a jej predpis je

$$F(x; T, \alpha) = 1 - \alpha e^{T \log(x)} \mathbf{1}, \tag{15}$$

kde $e^{T \log(x)}$ predstavuje maticovú exponenciálu, $\mathbf{1}$ je vektor jednotiek rozmeru $(p \times 1)$, T znázorňuje $(p \times p)$ maticu, pre ktorú platí, že jej nediagonálne prvky sú nezáporné a je záporne definitná. Súčet prvkov vektora α, ktorý má rozmery $(1 \times p)$ musí dávať v súčte 1. Následne môžeme definovať hustotu rozdelenia ako

$$f(x; T, \alpha) = \frac{1}{x} \alpha e^{T \log(x)} t, \quad x \geq 1, t = -T\mathbf{1}. \tag{16}$$

Využitím grey wolf algoritmu je naším cieľom nájsť parametre daného rozdelenia, a to maticu T a vektor α aby sme dokázali čo najlepšie popísať naše dáta.

Optimalizáciou v článku autorom vyšli parametre pre dané rozdelenie

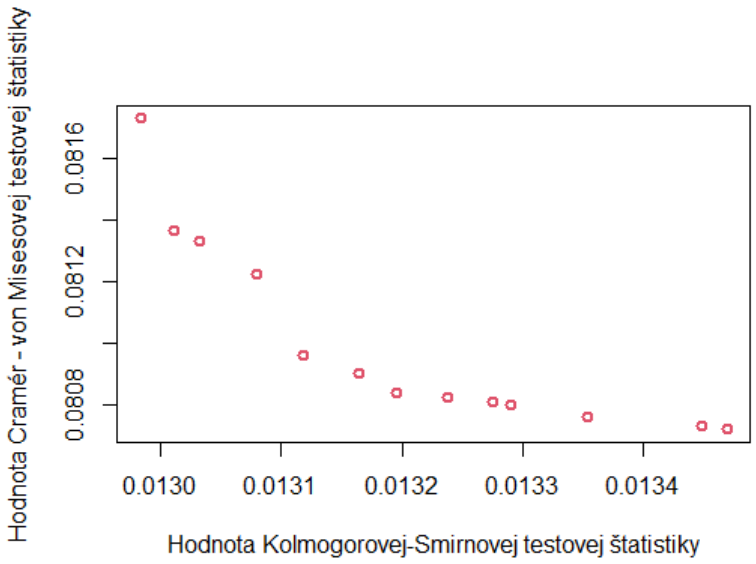
$$\hat{T} = \begin{bmatrix} -4.000 & 3.564 \\ 0.267 & -1.813 \end{bmatrix} \hat{\alpha} = \begin{bmatrix} 0.622 & 0.378 \end{bmatrix}. \tag{17}$$

Skutočnosť, že vektor α má dávať v súčte 1 a v našom prípade α je dvojrozmerný vektor, nám umožňuje zmenšiť náš optimalizačný problém len na 5 parametrov a následne druhý parameter vektora α sa dopočíta ako rozdiel $1 - \alpha(1)$, kde $\alpha(1)$ značí prvú zložku vektora. Ohraničenia pre parametre v matici sme zvolili tak, aby v intervale pokryli získané výsledky autormi článku a aby sme zabezpečili zápornosť diagonálnych členov a nezápornosť mimodiagonálnych členov. Kvôli zápornej definitnosti požadujeme záporné vlastné čísla matice T. Diagonálne prvky matice sme ohraničili intervalom $[-5, -0.01]$, mimodiagonálne prvky intervalom $[0.01, 5]$. Keďže súčet zložiek vektora α má dávať v súčte jednotku, parameter $\alpha(1)$ sme ohraničili najväčším možným intervalom $[0, 1]$.

Veľkosť populácie vlkov sme testovali s rôznymi parametrami, no pri veľkostiach 10,100 a 1000 sme nedostávali postačujúco dobré výsledky (v porovnaní s (17)). Veľkosť populácie sme nakoniec zvolili na hodnotu $greyWolves\_num = 10000$.

Za veľkú výhodu pokladáme skutočnosť, že v článku sa nachádza výsledný odhad parametrov (17), ku ktorému došli autori článku inou optimalizáciou. My tieto výsledky budeme brať ako dobrú príležitosť pre porovnanie dosiahnutých výsledkov naším algoritmom.

Ako výsledok sme dostali 13 Pareto optimálnych riešení. Hodnoty účelových funkcií pre dané riešenia znázorňujeme na Obr.7. Dá sa vidieť, že hodnoty sa líšia vo veľmi malých číslach. Ukážeme, že aj hodnoty parametrov sa líšia minimálne.



Obr. 7: Hodnoty účelových funkcií pre získané riešenia LogPH rozdelenia

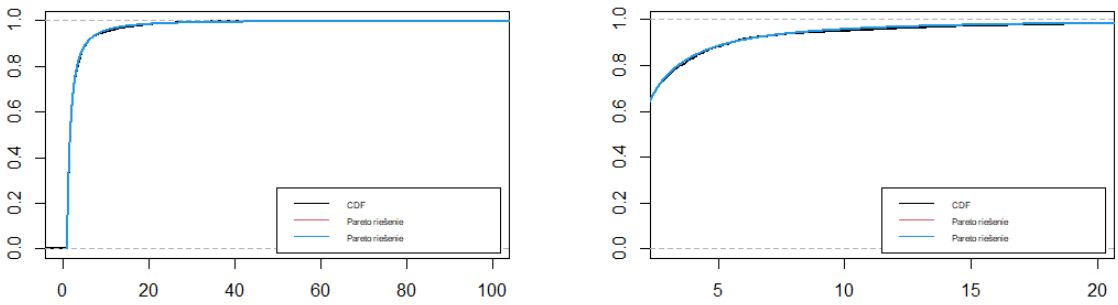
V Tabuľke 3 zaznamenávame vyčíslené testové štatistiky pre parametre z článku a zároveň najvyššiu a najnižšiu dosiahnutú K-S testovú štatistiku. Môžeme si všimnúť, že hodnoty testových štatistík sú v každom prípade menšie ako hodnoty testových štatistík pre výsledné parametre získané v článku. Všimnime si, že obe hodnoty testových štatistík sú nižšie, ako v prípade Burrovho rozdelenia.

Parametre	**K-S Testová štatistika**	**C-vM Testová štatistika**
Z článku [2]	0.01679	0.10526
Pareto riešenie (MGWO)	0.01298	0.08072
Pareto riešenie (MGWO)	0.01347	0.08173

Tabuľka 3: Vyčíslené testové štatistiky pre získané Pareto optimálne riešenia

Na Obr.8 (aj s detailným priblížením) vidíme kumulatívnu distribučnú funkciu (čiernou farbou) a distribučné funkcie $LogPH$ rozdelenia, pre ktoré sme získali parametre $MGWO$ (Multikriteriálnym grey wolf optimalizačným algoritmom). Červnou farbou je

znázornená distribučná funkcia pre parametre, ktoré prislúchajú najnižšej hodnote K-S testovej štatistiky a Modrou farbou je znázornená distribučná funkcia pre parametre, ktoré prislúchajú najvyššej hodnote K-S testovej štatistiky. Parametre sú natoľko podobné, že sa dané krivky prekrývajú. Z pozorovaných výsledkov usudzujeme, že sa nám podarilo nájsť také parametre, vďaka ktorým sa vskutku čo najlepšie približujeme ku empirickej kumulatívnej distribučnej funkcii.

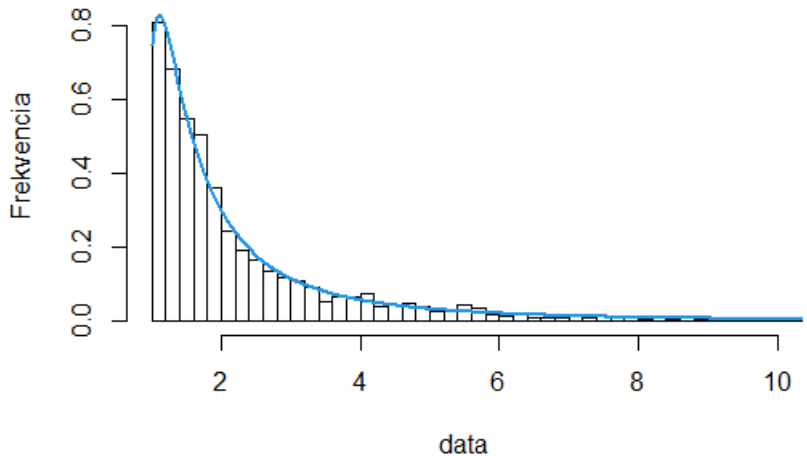


Obr. 8: Porovnanie distribučných funkcií LogPH rozdelenia pre Dánske dáta

Obr.9 znázorňuje, ako dobre hustota popisuje histogram dát. Krivka hustoty je znázornená pre parametre prislúchajúce najväčšej a najmenšej K-S testovej štatistike. Opäť vidíme, ako sú krivky natoľko podobné, že sa na obrázku prekrývajú.

2.2 S&P 500

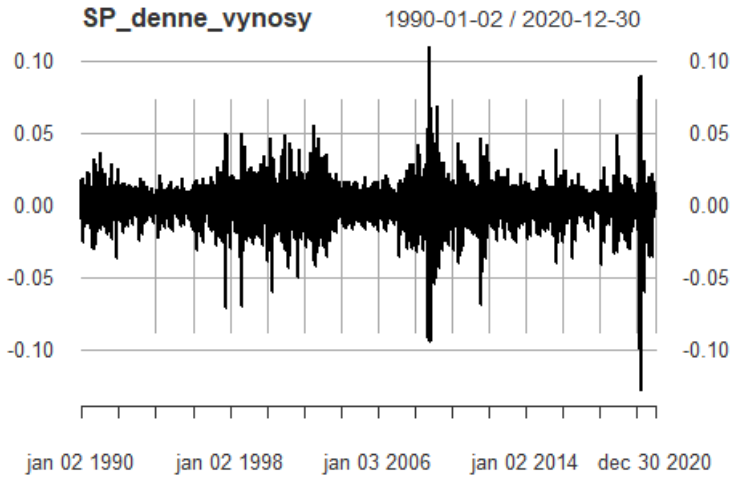
Asi najznámejší model pre oceňovanie finančných produktov, Black-Scholesov model pre modelovanie cien akcií, predpokladá normalitu dát. V tejto časti práce sa pokúsime odhadnúť parametre normálneho rozdelenia pre výnosy trhového indexu $S\&P$ 500. V týchto dvoch častiach (kedy sa budeme venovať rozdeleniu $S\&P$ 500) vychádzame z [4] a [9], kde sa autori snažili nájsť čo najlepšie alternatívne rozdelenia k normálnemu, aby čo najlepšie popisovali dáta trhového indexu $S\&P$ 500. V [4] ako alternatívy k normálnemu rozdeleniu zvolili zmes Gaussových rozdelení, zovšeobecnené logF rozdelenie a zovšeobecnené hyperbolické rozdelenie. Autori [9] si vybrali vyše desať rozdelení, ako napríklad normálne, chauchy, laplasové, studentovo, zošikmené verzie týchto rozdelení, Hyperbolické, variance gama či zovšeobecnené Hyperbolické rozdelenie. Porovnaním výsledkov v



Obr. 9: Porovnanie hustôt s histogramom LogPH rozdelenia pre Dánske dáta

závere konštatujú, že najlepšie rozdelenie z vybraných, ktoré by čo najlepšie fitovalo dáta trhového indexu $S\&P$ 500 boli hyperbolické, variance gama a zovšeobecnené hyperbolické rozdelenie. Táto trojica rozdelení bola vybraná aj pre trhový index $KOSPI$ 200.

Dáta, ktoré chceme nafitovať, predstavujú denné výnosy $S\&P$ 500 od začiatku roka 1990 po koniec roka 2020, ktoré sme v jazyku $R - studio$ získali funkciou *getSymbols* z balíka *quantmod* [33]. Získané denné výnosy sú znázornené v grafe na Obr.10.



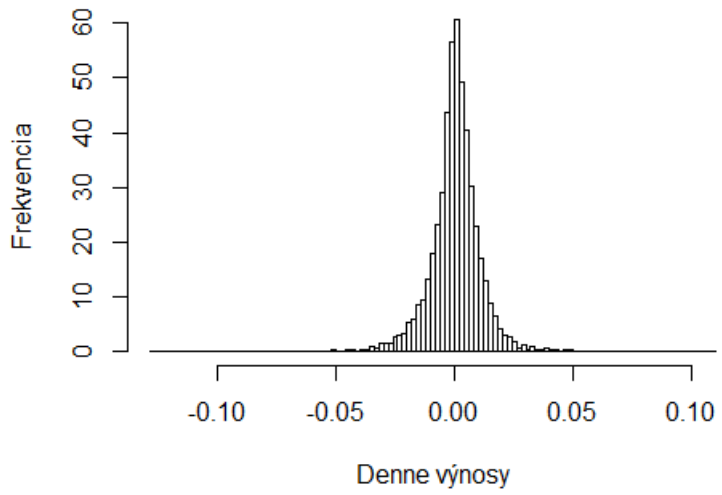
Obr. 10: Denné výnosy trhového indexu S&P 500

V Tabuľke 4 uvádzame základné štatistiky nášho dátového súboru.

Minimum	1. Kvantíl	Median	Priemer	3. Kvantíl	Maximum
-0.1276522	-0.0044005	0.0005742	0.0003018	0.0056259	0.1095720

Tabuľka 4: Základné štatistiky pre denné výnosy trhového indexu $S\&P$ 500

Taktiež vykreslíme histogram denných výnosov Obr. 11 z ktorého vidíme symetriu dát. Môže sa zdať, že dáta majú normálne rozdelenie, ktoré sa v nasledujúcej časti pokusíme nakalibrovať pre daný dátový súbor.

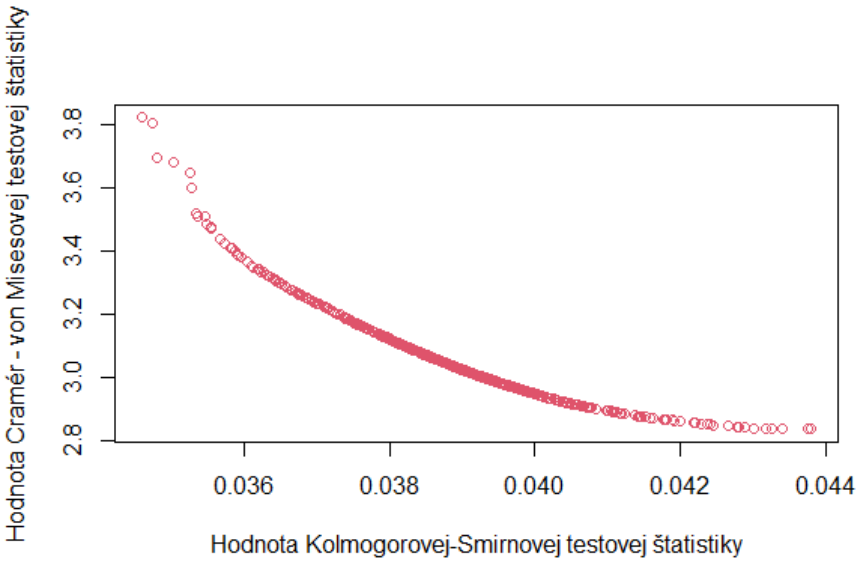


Obr. 11: Histogram denných výnosov S&P 500

2.2.1 Normálne rozdelenie

Distribučnú funkciu normálneho rozdelenia v matematickom softvéri $R - studio$ získame príkazom *pnorm*. Ohraničenia pre dané parametre sme zvolili následovne:
$\mu \in [\min(vynosy), \max(vynosy)]$, $\sigma \in [0, 1]$ a $greyWolves\_num = 500$. Ostatné parametre pre algoritmus ostávajú nastavené ako v predchádzajúcom prípade.

Pri takto zadaných parametroch získavame 446 Pareto optimálnych riešení, ktoré môžeme vidieť na Obr.12. Os x znázorňuje hodnotu prvej účelovej funkcie a na osi y je znázornená hodnota druhej účelovej funkcie pre získané parametre. Na obrázku sa dá pozorovať, ako naše nedominované riešenia spolu vytvárajú Pareto front.



Obr. 12: Pareto optimálne riešenia normálneho rozdelenia pre S&P 500

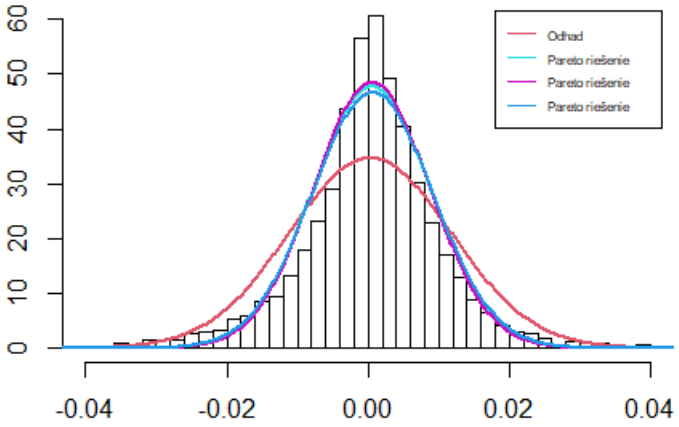
Na niekoľko príkladoch z týchto Pareto optimálnych riešení sme sa pozreli, ako dobre bude hustota s výslednými parametrami kopírovať histogram (Obr.11).

Na Obr.13 sme si vzali také riešenia pre parametre, kde je stredná hodnota z dosiahnutých riešení najväčšia $\mu = 0.0006525140$ (krivka modrej farby), najmenšia $\mu = 0.0005067258$ (krivka tyrkysovej farby) a medzi extrémami (krivka ružovej farby) získaných riešení $\mu = 0.0005427712$.

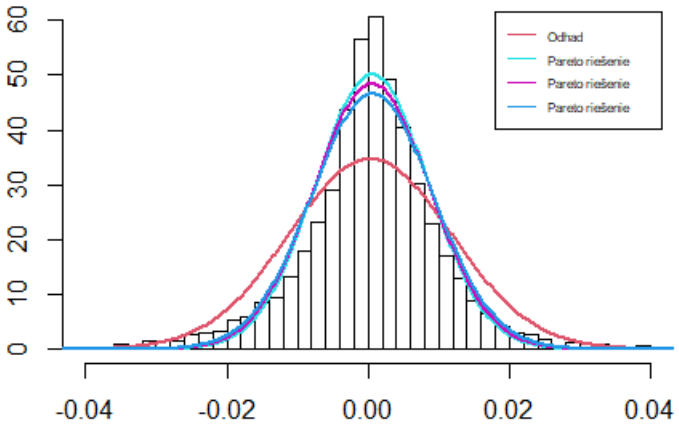
Obr. 14 popisuje hustoty pre riešenia, kde štandardná odchýlka je najväčšia $\sigma = 0.008564643$ (krivka modrej farby), najmenšia $\sigma = 0.007965095$ (krivka tyrkysovej farby) a niekde medzi týmito krajmi (krivka ružovej farby) z dosiahnutých výsledkov ($\sigma = 0.008247106$).

Obr. 15 znázorňuje hustotu pre parametre, kde testové štatistiky boli najextremnejšie (najvyššia KS štatistika (0.0438) - krivka modrej farby, najnižšia KS štatistika (0.0346) - krivka tyrkysovej farby) a keď sa nachádzali „niekde v strede“ nájdených riešení (krivka ružovej farby) - hodnota KS testovej štatistiky 0.0387.

Červenou farbou vidíme vykreslenú hustotu s parametrami, pre ktoré platí, že μ je priemer dát a σ je štandardnou odchýlkou. Na výsledných grafoch pozorujeme, že dáta odhadnuté algoritmom omnoho lepšie popisujú dáta, ako keby sme za strednú hodnotu a disperziu zvolili priemer a štandardnú odchýlku. Avšak, ak by sme zvolili nami nájdené parametre pre normálne rozdelenie, stále vidíme nedostatky. Stále nedokážeme dostatočne



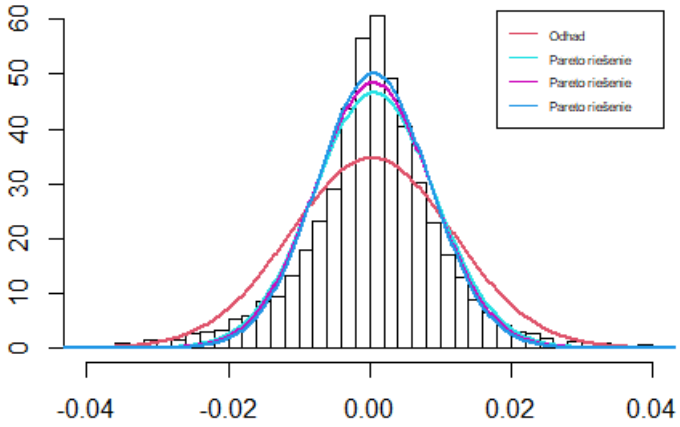
Obr. 13: Porovnanie Pareto optimálnych riešení normálneho rozdelenia pre S&P 500 pre najväčšiu, najmenšiu a „prostrednú“ nakalibrovanú strednú hodnotu



Obr. 14: Porovnanie Pareto optimálnych riešení normálneho rozdelenia pre S&P 500 pre najväčšiu, najmenšiu a „prostrednú“ nakalibrovanú štandardnú odchýlku

zachytiť vrchol histogramu dát a taktiež dostatočne pokryť chvosty. Už voľným okom z pozorovaných výsledkov môžeme povedať, že dané rozdelenie nepopisuje dáta dostatočne.

Keďže sú naše účelové funkcie vystavané tak, aby sme sa čo najviac priblížili empirickej



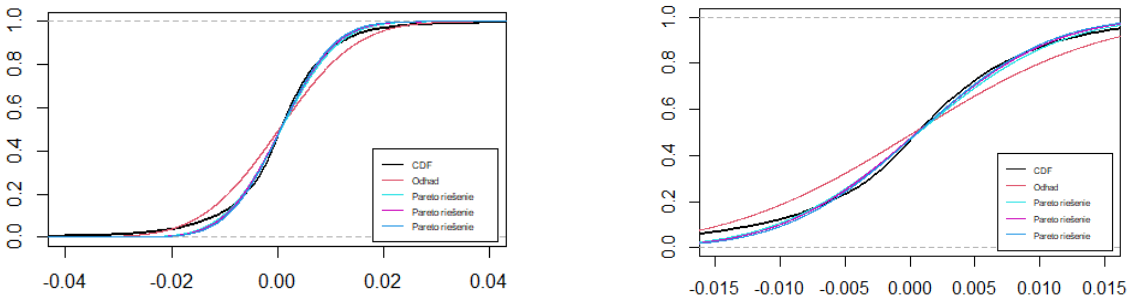
Obr. 15: Porovnanie Pareto optimálnych riešení normálneho rozdelenia pre S&P 500 pre najväčšiu, najmenšiu a „prostrednú“ výslednú testovú štatistiku

kumulatívnej distribučnej funkcii, tak sa pozrieme, ako blízko k empirickej distribučnej funkcii sme s našimi nakalibrovanými funkciami. Na Obr.16 vidíme, že distribučné funkcie získané parametrami $MGWO$ algoritmom omnoho lepšie kopírujú empirickú kumulatívnu distribučnú funkciu, ako distribučná funkcia získaná odhadom parametrov ako priemer a štandardná odchýlka z dát. Prikládame aj obrázok, na ktorom je zúžený interval x-ovej osi, kvôli detailnejšiemu pohľadu na dané nakalibrované krivky. Vybrané krivky prislúchajú parametrom, pre ktoré Kolmogorovova - Smirnovova testová štatistika nadobúda najväčšiu hodnotu, najmenšiu hodnotu a hodnotu testovej štatistiky medzi extrémami získaných výsledkov.

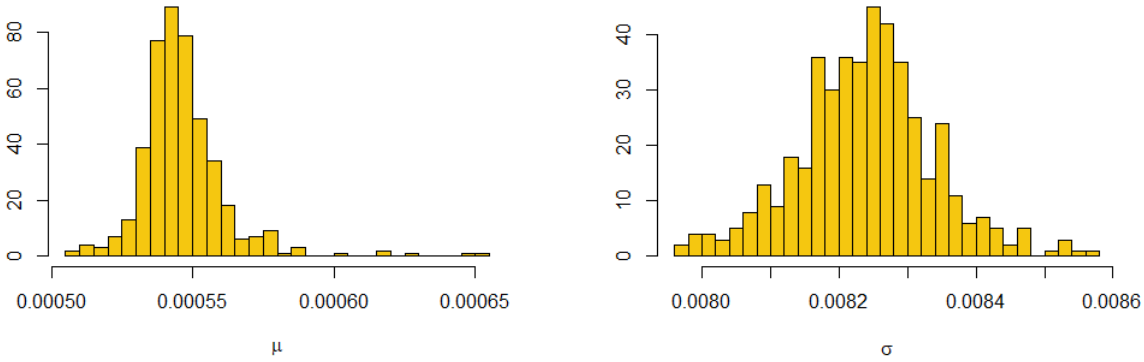
Histogramy na obrázku 17 nám znázorňujú hodnoty parametrov μ, σ a ich frekvenciu, ktoré sme získali v naších Pareto optimálnych riešeniach.

V grafe na obrázku 18 vidíme vzájomné porovnanie medzi získanými výsledkami strednej hodnoty μ a štandardnej odchýlky σ normálneho rozdelenia.

Pozorujeme, že od hodnoty parametra $\sigma = 0.0084$ sa zmení trend vzájomných výsledných parametrov na rastúci. Pozrieme bližšie, ktorým hodnotám účelových funkcií prislúchajú jednotlivé body. Na Obr.19 dané outliere zaznačujeme modrou farbou. Vidíme, že práve pre tieto parametre K-S testová štatistika nadobúda najnižšie hodnoty.



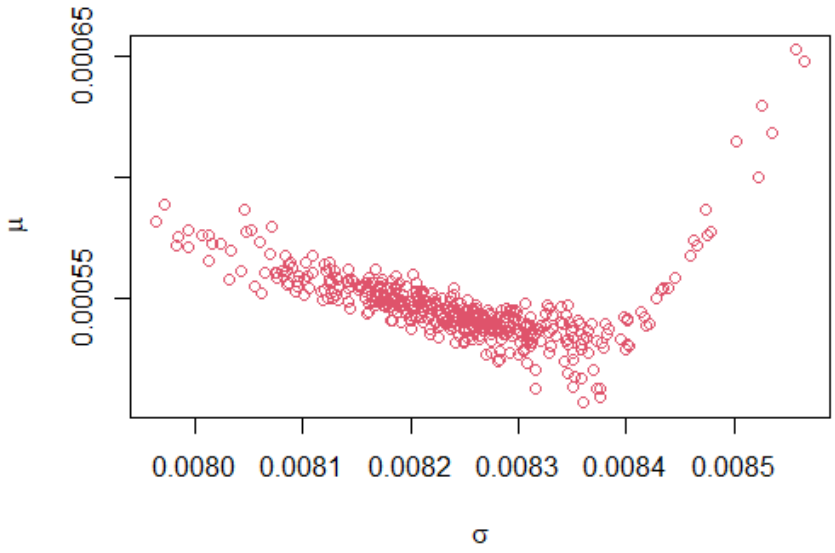
Obr. 16: Distribučné funkcie normálneho rozdelenia pre odhadované riešenie (vypočítané priemerom a varianciou) a Pareto optimálne riešenie získané algoritmom



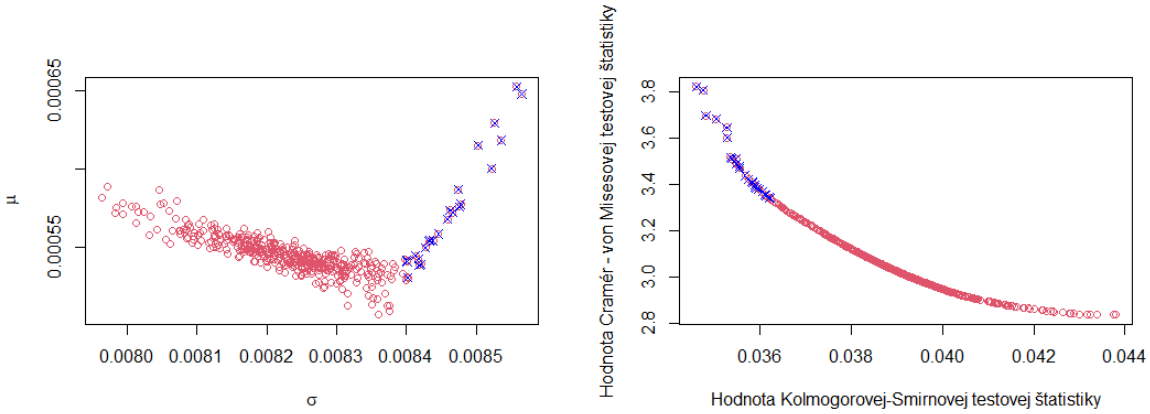
Obr. 17: Frekvencie získaných Pareto optimálnych riešení

2.2.2 Variance gama rozdelenie

Vychádzajúc z [9] sa dá vidieť, že autorom, ktorí sa snažili nájsť čo najlepšie rozdelenie pre trhový index $S\&P$ 500 sa už na začiatku normálne rozdelenie nepozdávalo. Ako sme spomínali vyššie, autori zhodnotili, že najlepšie z vybraných rozdelení ktoré by popisovali denné výnosy $S\&P$ 500 sa javia byť Hyporbolické, Variance gama a zovšeobecnené Hyperbolické rozdelenie. My sme sa na základe daného článku rozhodli, že sa pokúsime nakalibrovať Variance gama rozdelenie, ktoré je definované štyrmi parametrami.



Obr. 18: Vzájomné porovnanie parametrov μ a σ Normálneho rozdelenia pre Pareto optimálne riešenia



Obr. 19: Výber outlierov: vzájomné porovnanie parametrov μ a σ Normálneho rozdelenia pre Pareto optimálne riešenia a následne porovnanie hodnôt testových štatistík

Prepis funkcie hustoty daného rozdelenia je nasledovný:

$$f(x;\mu,\sigma,\alpha,\lambda) = \sqrt{\frac{\pi}{2}}\frac{\lambda^{1/\lambda}\exp\{((\alpha(x-\mu))/\sigma^2)\}}{\sigma\Gamma(1/\lambda)}. \\ \left(\frac{|x-\mu|}{\sqrt{\alpha^2+(2\sigma^2/\lambda)}}\right)^{1/\lambda-1/2}.K_{1/\lambda-1/2}\left(\frac{(x-\mu)\sqrt{\alpha^2+(2\sigma^2/\lambda)}}{\sigma^2}\right) \quad (18)$$

kde $\mu, \sigma, \alpha, \lambda$ predstavujú parametre. $\lambda > 0$, $\Gamma(.)$ je Gama funkcia a K_λ symbolizuje

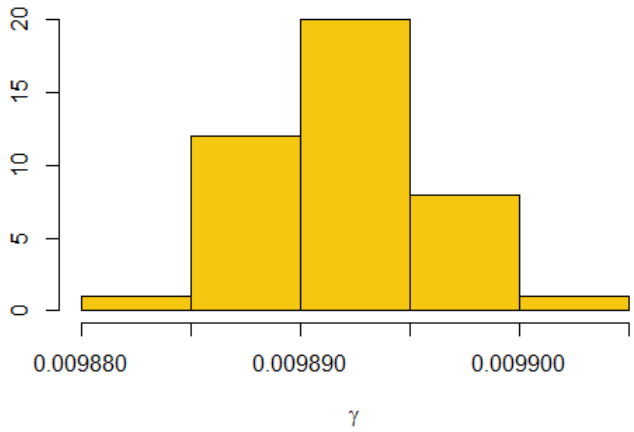
modifikovanú Basselovú funkciu s indexom λ. V $R - studiu$ distribučnú funkciu získame implementovanou funkciou *pvg*, kde ako vstupné parametre potrebujeme zadať naše dáta a následne odhadované parametre. Treba dať pozor, keďže daná funkcia môže byť definovaná vo viacerých zdrojoch inak. Daný problém sa dá v softvéri ošetriť spätnou transformáciou parametrov pomocou funkcie *vgChangePars*. Implementovaná hustota rozdelenia *dvg* sa nezhoduje s funkciou, ktorú sme definovali vyššie a má iný prepis (používa inú transformáciu premenných). V [34] taktiež kalibrovali Variance gama rozdelenie pre výnosy akcií $S\&P$ 500. V tomto článku kalibrovali rozdelenie dané prepisom hustoty a distribúcie rovnakým, aký využíva pri výpočte implementovaná funkcia pre Variance gama rozdelenie v $R - studiu$. Preto pri určovaní intervalov prípustnosti pre riešenia sme vychádzali z dosiahnutých výsledkov pre parametre, ktoré vyšli pri kalibrovaní $S\&P$ 500 v [34]. Určiť dobrý interval sa javí byť kľúčové, pretože s príliš veľkým intervalom prípustnosti sa nám nepodarilo nájsť dostatočne dobré riešenia.

Keďže náš algoritmus funguje sčasti náhodnou voľbou parametrov, ľahko sa môže stať, že funkcia *pvg* (resp. *dvg*), ktorá na výpočet distribúcie a hustoty využíva funkciu *optim*, nebude schopná nájsť optimálne riešenie pre výpočet distribučnej funkcie a hustoty. Aj na túto skutočnosť je potrebné myslieť pri programovaní a hľadaní optima.

Algoritmom sme našli 42 Pareto optimálnych riešení. Hodnoty parametrov pre získané riešenia uvádzame v Tab.5. Keďže pri kalibrácii sme použili funkcie, s transformovanými premennými (vzhľadom na pôvodné premenné μ, σ, α a λ), parametre zapíšeme ako $\beta, \gamma, \theta, \nu$. Parametre β, θ a ν vyšli vo všetkých prípadoch rovnaké. Histogram pre parameter γ uvádzame na Obr.20. Vidíme, že parameter γ vychádza rôzne, no vo veľmi úzkom intervale. Z tohoto usudzujeme, že získané výsledné parametre nie sú dostatočne rôzne na to, aby sme ich považovali za rozdielne. To potvrdíme aj vykreslením distribučnej funkcie a hustoty pre parametre prislúchajúce najextrémnejším hodnotám testovej štatistiky.

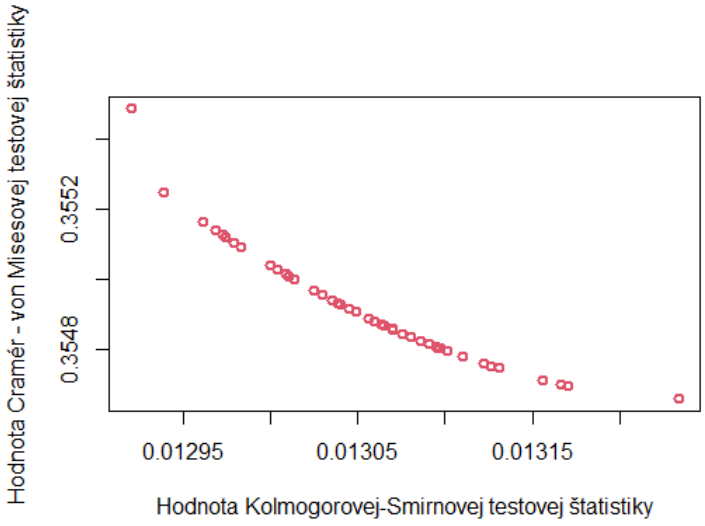
β	θ	ν
0.000517	-0.0000755	0.844

Tabuľka 5: Porovnanie výsledných parametrov Variance gama rozdelenia pre získané Pareto optimálne riešenia



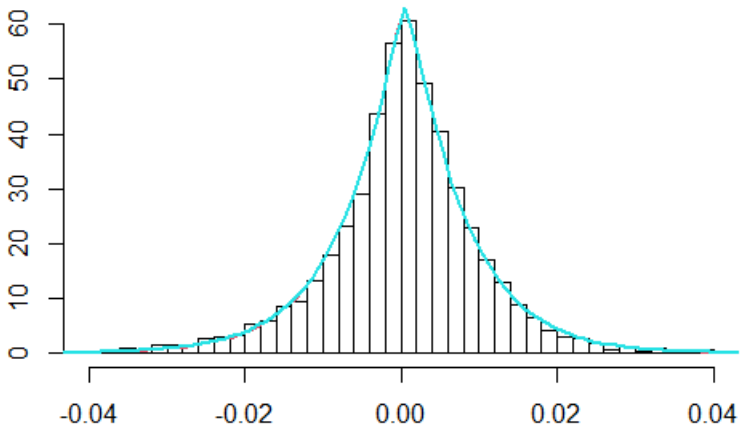
Obr. 20: Histogram parametra γ Variance gama rozdelenia pre S&P 500

Obr. 21 znázorňuje získané hodnoty účelových funkcií vykreslené v grafe. Na osi x je hodnota K-S testovej štatistiky a na osi y sa nachádza hodnota C-vM testovej štatistiky. Zároveň tu pozorujeme už spomínanú vlastnosť týchto riešení. So snížujúcou sa hodnotou jednej optimalizačnej funkcie hodnota druhej optimalizačnej funkcie narastá.



Obr. 21: Získané Pareto optimálne riešenia Variance gama rozdelenia pre S&P 500

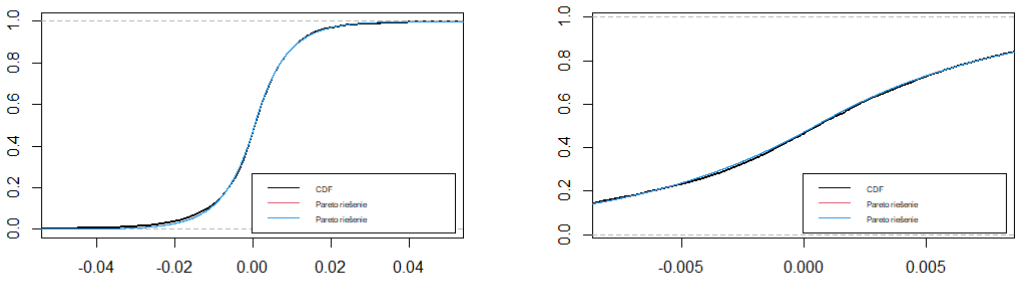
Na Obr. 22 vidíme ako hustota daného rozdelenia kopíruje histogram dát. Červenou farbou je znázornená krivka popisujúca hustotu pre najnižšiu hodnotu K-S testovej štatistiky, a naopak, modrou farbou krivka určená parametrami, ktoré prislúchajú najvyššej



Obr. 22: Hustota Variance gama rozdelenia pre získané riešenia

hodnote K-S testovej štatistiky. Vidíme, že dané krivky sú v podstate identické. Je nespochybniteľné, že Variance gama rozdelenie omnoho lepšie kopíruje histogram trhového indexu $S\&P$ 500 ako normálne rozdelenie.

Môžeme sa pozrieť aj na porovnanie kumulatívnej distribučnej funkcie a Variance gama distribučnej funkcie s odhadnutými parametrami (Obr. 23). Farebné krivky (červená a modrá) predstavujú distribučné funkcie Variance gama rozdelenia a čierna krivka znázorňuje kumulatívnu distribučnú funkciu. Krivky distribučných funkcií rozdelenia kopírujú kumulatívnu distribučnú funkciu omnoho presnejšie ako v prípade Normálneho rozdelenia. Opäť vidíme, že krivky popísané parametrami, ktoré patria najextrémnejším prípadom K-S testovej štatistiky sú takmer zhodné.



Obr. 23: Distribučná funkcia pre Variance gama rozdelenie

3 Optimalizácia portfólia

Využitím grey wolf optimalizáčneho algoritmu sme sa v tejto časti práce pokúsili odhadnúť výšky váh jednotlivých akcií pre zostavenie optimálneho portfólia. Najprv sa pozrieme na Markowitzov model, kde budeme minimalizovať riziko odhadnuté varianciou portfólia a zároveň maximalizovať výnos portfólia. Získané výsledky porovnáme s efektívnou hranicou Markowitzového portfólia, ktorú získame ako výsledok kvadratickej optimalizácie.

V druhej časti riziko portfólia nahradíme Mean-Gini kritériom ktoré budeme chcieť minimalizovať a zároveň budeme maximalizovať výnos portfólia ako v prvej časti optimalizácie portfólia.

3.1 Markowitzova optimalizácia portfólia

Vychádzajúc z teórie Markowitzového portfólia optimalizačná funkcia, pre ktorú budeme hľadať minimum naším genetickým algoritmom je definovaná nasledovne [25]

$$\min_w w^T \Sigma w, \tag{19}$$

kde w predstavuje vektor váh rozmeru $N \times 1$ a matica Σ rozmeru $N \times N$ predstavuje kovariančnú maticu denných výnosov vybraných aktív. Súčasne chceme maximalizovať výnos daných akcií

$$\max_w w^T \mu, \tag{20}$$

čo predstavuje druhú optimalizačnú funkciu naším algoritmom. Vektor μ rozmeru $N \times 1$ hovorí o priemernom ročnom výnose aktív, resp. i-tá zložka vektora μ hovorí o priemernom

ročnom výnose i-tého aktíva. Váhy pre jednotlivé aktíva ohraničíme len na nezáporné hodnoty (nepočítame s krátkymi pozíciami) $w_i \geq 0$ a zároveň nastavíme podmienku, že celý investovaný kapitál rozdelíme medzi dané aktíva ($\sum_1^N w_i = 1$).

3.2 Mean-Gini model

Ako alternatívny odel k Markowitzovmu modelu sme zvolili Mean-Gini model vychádzajúc z [22]. Hlavnou myšlienkou nášho nového kritéria je, pozrieť sa na absolútne odchýlky výnosov akcií medzi všetkými časovými bodmi. Predpokladajme, že máme N akcií, pre ktoré máme časové rady historických výnosov. Ako w_j označíme podiel investovaného kapitálu do j-tej akcie, $r_{j,t}$ predstavuje výnos j-tého aktíva v periode t, kde $t = 1, ..., T$ a T predstavuje konečný počet časových periód. V našej aplikácii budeme počítať s týždennými periódami.

Dvojkriterálny Mean-Gini model vyzerá následovne

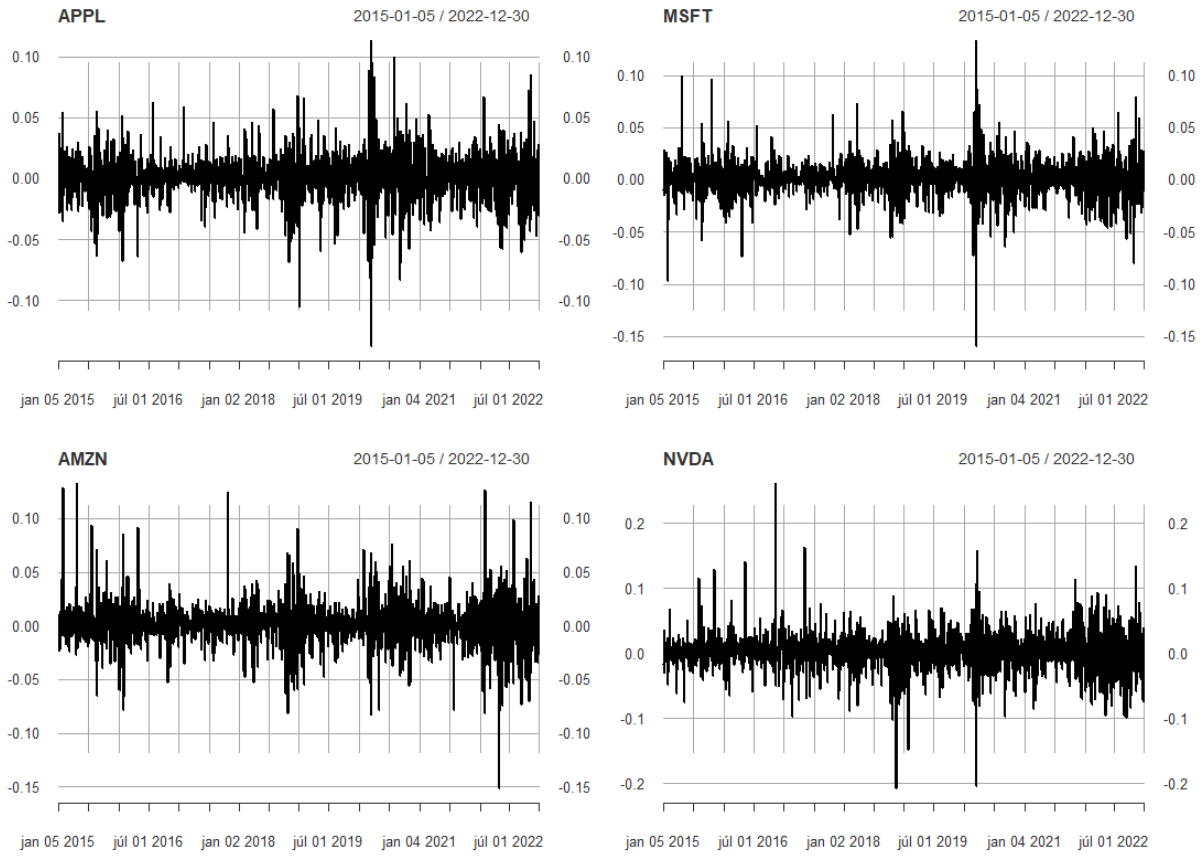
$$\begin{aligned} \min_{w} \quad & G = \frac{1}{T(T-1)} \sum_{t=1}^{T-1} \sum_{\bar{t}>t}^{T} |\sum_{j=1}^{N} w_j (r_{j,t} - r_{j,\bar{t}})| \\ \max_{w} \quad & w^T \mu \\ & w_i \geq 0 \qquad i = 1, ..., N \\ & \sum_{1}^{N} w_i = 1. \end{aligned} \tag{21}$$

Ohraničenia pre premenné ostávajú rovnaké ako v predchádzajúcom prípade.

3.3 Výnosy akcií Applu, Microsoftu, Amazonu a Nvidii

Optimalizáciu portfólia pre vyššie spomenuté kritéria sme sa rozhodli aplikovať pre 4 najvýnosnejšie akcie z trhového indexu $S\&P$ 500 (podľa [32]), ktorými sú Apple Inc., Microsoft Corporation, Amazon.com Inc. a NVIDIA Corporation. Na Obr.24 znázorňujeme priebeh daných výnosov v časovom rozpätí 01.01.2015 do 01.01.2023.

Zároveň v Tab.6 uvádzame základné štatistiky pre vybrané dátové súbory.



Obr. 24: Priebeh výnosov vybraných akcií

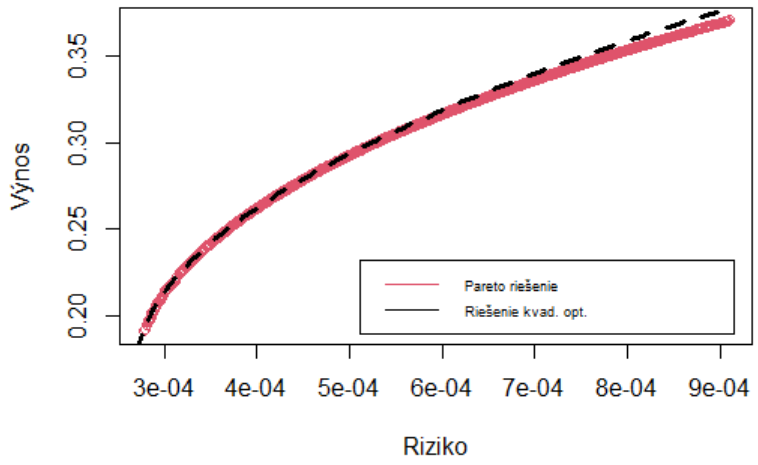
Akcia	Minimum	1. Kvantíl	Median	Priemer	3. Kvantíl	Maximum
APPL	-0.1377079	-0.0075969	0.0007649	0.0008267	0.0103021	0.1131574
MSFT	-0.1594536	-0.0066858	0.0008444	0.0008797	0.0097967	0.1329292
AMZN	-0.1513979	-0.0085804	0.0011677	0.0008419	0.0108334	0.1321778
NVDA	-0.207712	-0.012043	0.002459	0.001693	0.016853	0.260876

Tabuľka 6: Základné štatistiky pre výnosy vybraných akcií

3.3.1 Použitie Markowitzovho modelu

Najprv sa pozrieme na optimálne váhy získané Markowitzovou optimalizáciou portfólia pre vybraný dátový súbor. Spolu s Pareto optimálnymi riešeniami (pre populáciu 500 vlkov a 10 iterácií) získavame 1338 Pareto optimálnych riešení. Na Obr.25 červenou farbou znázorňujeme hodnoty účelových funkcií (výnos a riziko portfólia) pre Pareto optimálne riešenia. Zároveň, čiernou farbou zakreslíme hodnotu optimálneho riešenia, ktoré získame

kvadratickou optimalizáciou, kde minimalizujeme riziko portfólia, pri fixne danom výnose. Vidíme, že takto získaná hranica optimálneho portfólia sa v celku zhoduje s nami nájdenými optimálnymi riešeniami. Výnos portfólia sa pohybuje od 0.19 do 0.37 a riziko od 0.0002775 po hodnotu 0.0009098.



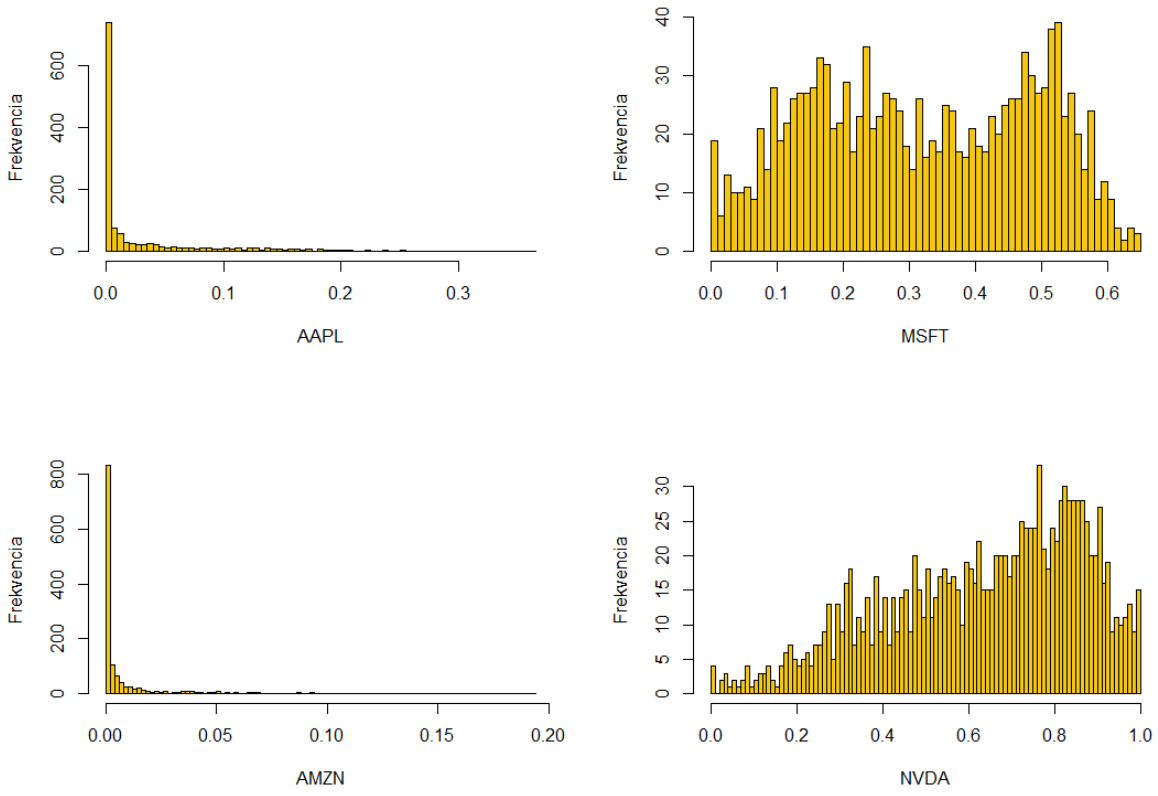
Obr. 25: Pareto riešenia (červenou farbou) a efektívna hranica (čiernou farbou) Markowitzovej optimalizácii portfólia

Následne udávame histogramy, ktoré znázorňujú početnosti váh získaných optimálnych riešení pre dané aktíva (Obr.26). Do akcií Applu a Amazonu by sme vo väčšine prípadov neinvestovali. Investície do Microsoftu a Invidie sú viacej diverzifikované. Ako investor by sme boli ochotní do akcií Microsoftu investovať od 0 do 65 percent investovaného kapitálu, pričom v priemere by sme do danej akcii investovali 31% z kapitálu. Najväčší obnos peňazí by sme zanechali pre Invidiu. Škála investovaného kapitálu sa pohybuje od 0 do 100 percent, no v priemere by sme do Invidie investovali až 63% z kapitálu

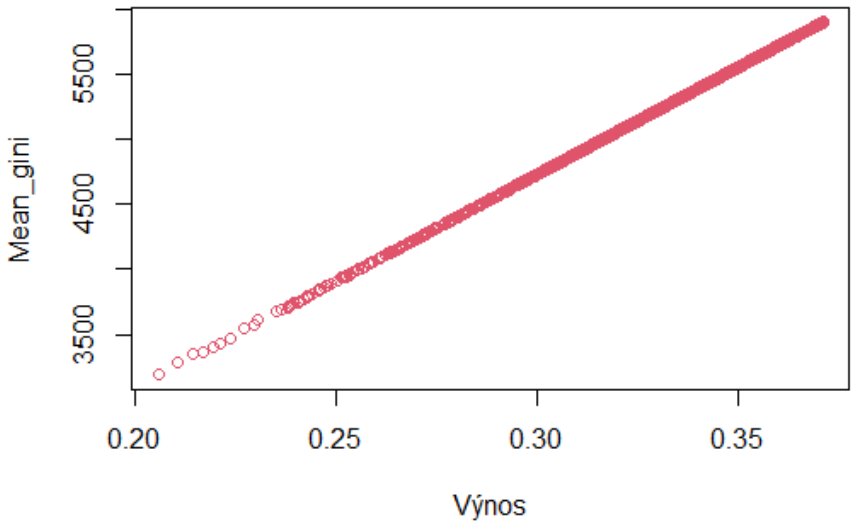
3.3.2 Použitie Mean-Gini modelu

Výmenou účelovej funkcie, popisujúcej riziko v Markowitzovom prístupe, za Mean-Gini kritérium, dostávame 3119 Pareto optimálnych riešení (Obr. 27), kde sa výnos pohybuje v rovnakom rozpätí ako v predošlej optimalizácii.

Z histogramov vidíme veľmi podobné rozloženie váh, ako pri optimalizácii Markowitzovho portfólia. Môžeme povedať, že do Applu a do Amazonu by sme neinvestovali.

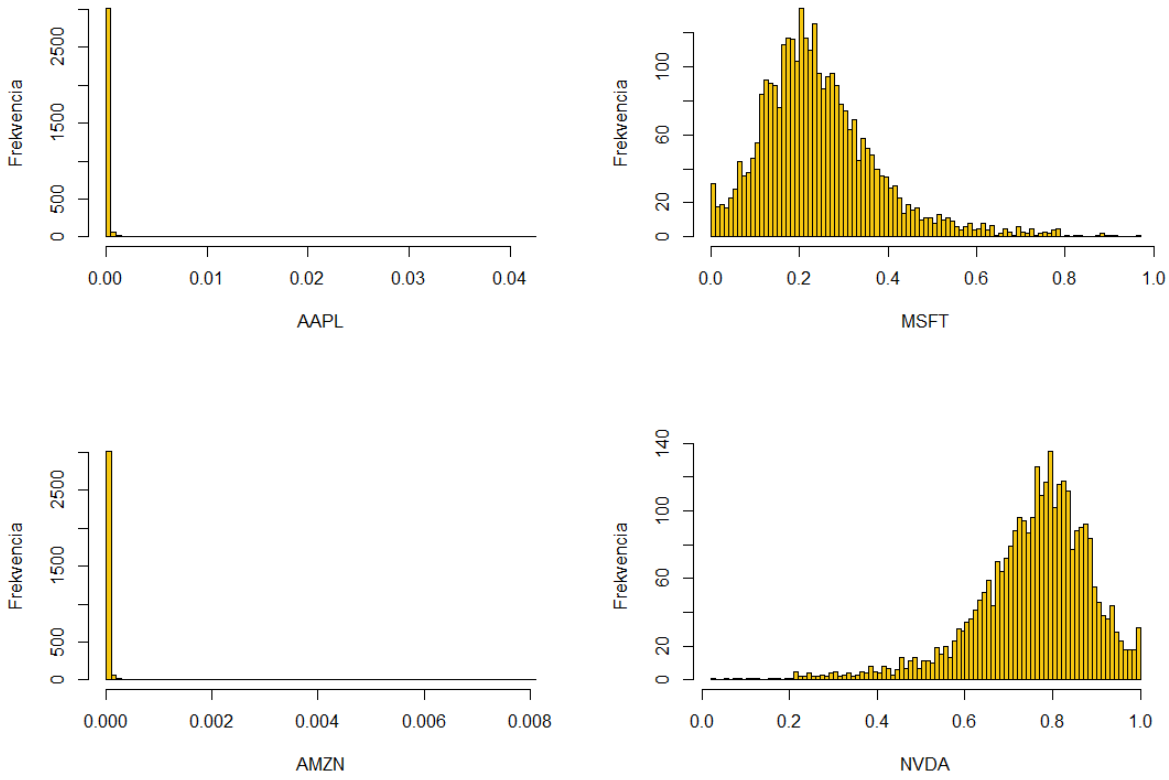


Obr. 26: Histogramy váh pre získané riešenia v Markowitzovej optimalizácii portfólia



Obr. 27: Pareto riešenia pre Mean-Gini kritérium

Výraznejší rozdiel pozorujeme pri Microsofte, kde výška investície sa pohybuje od 0 do 60 percent kapitálu a priemerná výška investície pre danú akciu je 25%. Najvyšší podiel z investície patrí Invidii. V priemere až 75% z kapitálu by išiel pre túto akciu (čo tvorí doplnok k akciám Microsoftu).



Obr. 28: Histogramy váh pre získané riešenia pre Mean-Gini model

4 Kalibrácia úrokových mier

S úrokovými mierami sa vo financiách stretávame pri diskontovaní, kde vstupujú úrokové miery pre zohľadnenie peňažných tokov v jednotlivých časových úsekoch. Na trhu nachádzame finančné deriváty, ktoré majú výplatu v závislosti od úrokových mier.

Existujú modely [25], ktoré definujú krátkodobé úrokové miery stochastickým procesom. Pridaním trhovej ceny rizika dokážeme namodelovať cenu bondu pre dlhodobé úrokové miery, a z toho exaktne vyjadriť aj úrokové miery s dlhodobou maturitou.

Vašíčkov model sa pokúšali nakalibrovať využitím dvojkriterálnej optimalizácie v [21] využitím Kalmanovho filtra. Taktiež v [6] autori podrobnejšie skúmajú robustnosť algoritmu pre odhadovanie parametrov pre európske a domáce úrokové miery, kde na konkrétnych príkladoch pre rôzne maturity odhadujú vhodné parametre pre popis daných úrokových mier.

Vychádzajúc z [6], [11] sa zameriame na model, ktorý sa snaží namodelovať domáce úrokové miery, ktoré majú tendenciu správať sa podľa úrokových mier inej krajiny (v našom prípade eurozóny).

Za domácu úrokovú mieru sme sa rozhodli úrokovú mieru v Českej republike. Keďže je Česká republika súčasťou Európskej únie, môžeme predpokladať, že úroková miera v tejto krajine bude ovplyvnená európskou úrokovou mierou. Pre namodelovanie európskej úrokovej miery sa použije jednofaktorový Vašíčkov model. Následne česká úroková miera bude modelovaná závisle od európskej úrokovej miery.

4.1 Model

Vychádzajúc z [6], [11] zadefinujeme model, parciálne diferenciálne rovnice pre výpočet parametrov európskych a domácich úrokových mier a odvodíme zjednodušený výpočet parametrov pre domáce úrokové miery pri fixnom parametri rýchlosti konvergencie b. Definujme maturitu bondu ako T. Cenu daného bondu $P(t,T)$ v čase t s úrokovou mierou $R(t,T)$ v čase t môžeme vyjadriť nasledovne

$$P(t,T) = e^{-R(t,T)(T-t)} \quad \text{resp.} \quad R(t,T) = \frac{-\log(P(t,T))}{(T-t)}, \tag{22}$$

kde $T-t$ (zvykneme označovať ako τ) značí dĺžku času do maturity.

Krátkodobá úroková miera je modelovaná stochastickým procesom

$$dr_e = c(d - r_e)dt + \sigma_e dw_e, \tag{23}$$

$$dr_d = (a + b(r_e - r_d))dt + \sigma_d dw_e, \tag{24}$$

kde dw_e a dw_d predstavujú Wienerove procesy, ktoré môžu byť korelované, preto označme ρ ako korelačný koeficient, resp. $\mathbf{E}(dw_e dw_d) = \rho dt$.

Parameter d je rovnovážny stav krátkodobej európskej úrokovej miery, parameter $c > 0$ popisuje rýchlosť konvergenncie, $\sigma_e > 0$ a $\sigma_d > 0$ predstavujú rozptyl náhodných prírastkov a r_e (resp.r_d) zodpovedá európskym (resp. domácim) krátkodobým úrokovým mieram. Predpokladáme, že domáce úrokové miery sú v rovnovážnom stave priťahované k európskym, rýchlosťou danou parametrom $b > 0$.

Danou trhovou cenou rizika λ_e a jednofaktorovým Vašíčkovým modelom nájdeme pomocou parciálnej diferenciálnej rovnice cenu europského bondu $P(r_e, \tau)$, pre ktorý platí

$$-\frac{\partial P}{\partial \tau} + (c(d - r_e) - \lambda_e \sigma_e)\frac{\partial P}{\partial r_e} + \frac{\sigma_e^2}{2}\frac{\partial^2 P}{\partial r_e^2} - r_e P = 0. \tag{25}$$

Daná parciálna rovnica musí spĺňať počiatočnú podmienku $P(r_e, \tau) = 1$. Riešenie parciálnej rovnice zapíšeme v tvare

$$P(r_e, \tau) = \exp\{F(\tau) - r_e G(\tau)\}, \tag{26}$$

kde

$$G(\tau) = \frac{1 - \exp(-c\tau)}{c} \tag{27}$$

$$F(\tau) = \frac{(G(\tau) - \tau)(c(cd - \lambda_e \sigma_e) - \sigma_e^2)}{c^2} - \frac{\sigma_e^2 (G(\tau)^2)}{4c}. \tag{28}$$

Spätnou transformáciou dokážeme vyjadriť model pre dlhodobé európske úrokové miery s danou maturitou τ.

Cena domáceho bondu $P(r_e, r_d, \tau)$ závislá od európskej a domácej úrokovej miery musí spĺňať parciálnu diferenciálnu rovnicu

$$\begin{aligned} -\frac{\partial P}{\partial \tau} + (a + b(r_e - r_d) - \lambda_d \sigma_d)\frac{\partial P}{\partial r_d} + (c(d - r_e) - \lambda_e \sigma_e)\frac{\partial P}{\partial r_e} \\ + \frac{\sigma_d^2}{2}\frac{\partial^2 P}{\partial r_d^2} + \frac{\sigma_e^2}{2}\frac{\partial^2 P}{\partial r_e^2} - \rho \sigma_d \sigma_e \frac{\partial^2 P}{\partial r_d \partial r_e} = 0. \end{aligned} \tag{29}$$

λ_d predstavuje trhovú cenu rizika a počiatočná podmienka pre parciálnu diferenciálnu rovnicu je daná $P(r_e, r_d, \tau) = 1$.

Použitím transformácie parametrov

$$\begin{aligned} a_1 &= a - \lambda_d \sigma_d, \\ a_2 &= -b, \\ b_1 &= cd - \lambda_e \sigma_e, \\ b_2 &= -c \end{aligned}$$

a za predpokladu, že $\rho = 0$ sa dá ukázať, že riešenie parciálnej diferenciálnej rovnice (29) môžeme zapísať v tvare

$$\log P(r_e, r_d, \tau) = a_1 F_1(\tau, a_2) + \sigma_d^2 F_2(\tau, a_2) + F_3(\tau, a2) - D(\tau, a_2) r_d - E(\tau, a_2) r_e, \quad (30)$$

kde funkcie D a E definujeme ako

$$\begin{aligned} D(\tau) &= \frac{-1 + e^{a_2 \tau}}{a_2}, \\ E(\tau) &= -\frac{1}{b_2} \frac{a_2(1 - e^{b_2 \tau}) - b_2(1 - e^{a_2 \tau})}{a_2 - b_2} \end{aligned}$$

a funkcie F_1, F_2 a F_3 sú definované nasledovne

$$\begin{aligned} F_1(\tau, a_2) &= -\int_0^\tau D(s)\, ds, \\ F_2(\tau, a_2) &= \frac{1}{2} \int_0^\tau D^2(s)\, ds, \\ F_3(\tau, a_2) &= -b_1 \int_0^\tau E(s)\, ds + \frac{\sigma_e^2}{2} \int_0^\tau E^2(s)\, ds. \end{aligned}$$

Parametre b_1, b_2 a σ_e^2 získame nakalibrovaním európskych úrokových mier podľa (26). Pre fixný parameter a_2 dokážeme nájsť optimálne parametre a_1 a σ_d^2 následovne.

Chceme nájsť minimum funkcie

$$\min \sum_{i=1}^{N} (LRD_i - modelLRD_i)^2, \quad (31)$$

kde LRD predstavuje dlhodobé domáce úrokové miery, $modelLRD$ predstavuje modelované úrokové miery podľa (30) a suma ide cez všetky pozorované a modelované dáta. Označíme $y = LRD - F_3 + Dr_d + Er_e$ a minimalizačnú funkciu môžeme zapísať v tvare

$$\min \sum_{i=1}^{N} (y_i - a_1 F_1 - \sigma_d^2 F_2)^2.$$

Derivovaním podľa premennej a_1 a σ_d^2 dostávame vzťahy: $2\sum_{i=1}^{N}(y_i - a_1F_1 - \sigma_d^2F_2)F_1 = 0$ a $2\sum_{i=1}^{N}(y_i - a_1F_1 - \sigma_d^2F_2)F_2 = 0$.

Keďže F_1 a F_2 sú konštanty, môžeme ich z rovníc vykrátiť. Takto sme získali jednu rovnicu pre 2 neznáme. Zvolením ľubovoľného parametra σ_d^2 vidíme, že parameter a_1 musí spĺňať

$$a_1 = \frac{\sum y_i}{NF_1} - \frac{\sigma_d^2F_2}{F_1}, \tag{32}$$

kde N je počet pozorovaní. Keďže funkcie F_1 a F_2 sú závislé iba od parametra a_2, pri fixne zvolenom parametri σ_d^2 dostávame parameter a_1 závislý iba od parametra a_2. Teda model pre výpočet dlhodobých úrokových mier (30) je závislý len od jediného parametra a_2.

4.2 Aplikácia pre európske a české úrokové miery

Pokúsime sa Vašičkovým modelom odhadnúť parametre pre dlhodobé európske úrokové miery, a následne ako domáce namodelujeme české úrokové miery, kde predpokladáme, že tieto úrokové miery sú závislé od európskych. Dátové súbory pre krátkodobé a dlhodobé úrokové miery sťahujeme z [30]. Časové rozpätie pre pozorované dáta je dané od 01.04.2000 do 31.12.2020. Stiahnuté dáta predstavujú mesačné rezy, čo znamená, že máme k dispozícii 249 pozorovaní.

Na úvod v Tab.7 uvádzame základné štatistiky pre krátkodobé európske a domáce úrokové miery (SRE, SRD) s maturitou 3 mesiace a zároveň pre dlhodobé úrokové miery (LRE, LRD) s maturitou 10 rokov. Vidíme, že európske úrokové miery pre dané pozorované obdobie nadobúdajú aj záporné hodnoty, kým domáce iba kladné hodnoty.

	Minimum	**1. Kvantíl**	**Median**	**Priemer**	**3. Kvantíl**	**Maximum**
SRE	-0.005381	-0.000536	0.010420	0.015624	0.029857	0.051131
SRD	0.00280	0.00490	0.01910	0.01956	0.02580	0.05570
LRE	-0.000915	0.013330	0.036893	0.030925	0.042567	0.055155
LRD	0.00250	0.01680	0.03590	0.03236	0.04530	0.07590

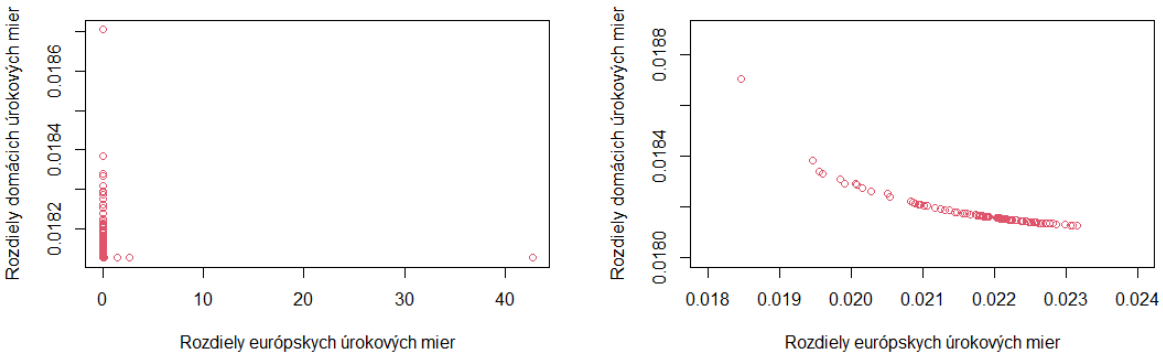
Tabuľka 7: Základné štatistiky pre európske a české úrokové miery

Optimalizujeme funkciu (31) vzhľadom na európske ale aj domáce úrokové miery, a

zvyšné parametre dopočítame vyššie spomínaným spôsobom. Veľkosť populácie v algoritme je nastavená na hodnotu 1000 a počet iterácií na hodnotu 10. Keďže parameter $c \geq 0$ vieme, že parameter $b_2 \leq 0$. Ohraničíme ho intervalom [-0.1,0] a σ_e^2 intervalom [0,0.1]. Parameter b_1 sme ohraničili širším intervalom [-0.5,0.5]. Takto získavame 95 Pareto optimálnych riešení (Obr. 29). V niekoľkých riešeniach pozorujeme výraznejší odklon optimalizačnej funkcie zodpovedajúcej európskym úrokovým mieram od väčšiny hodnôt účelových funkcií. Riešenia prislúchajúce týmto outlierom zanedbávame - pokladáme ich za neefektívne. Okrem týchto outlierov hodnoty optimalizačných funkcií sa pohybujú vo veľmi úzkom intervale. Pre prehľad udávame aj základné štatistiky pre získané hodnoty optimalizačných funkcií, spomedzi ktorých sme vyhodili spomínaných outlierov.

Optim. funkcia	**Minimum**	**1. Kvantíl**	**Priemer**	**3. Kvantíl**	**Maximum**
$(LRE - LRE_{Model})^2$	0.01846	0.02104	0.02170	0.02237	0.02315
$(LRD - LRD_{Model})^2$	0.01813	0.01815	0.01819	0.01821	0.01871

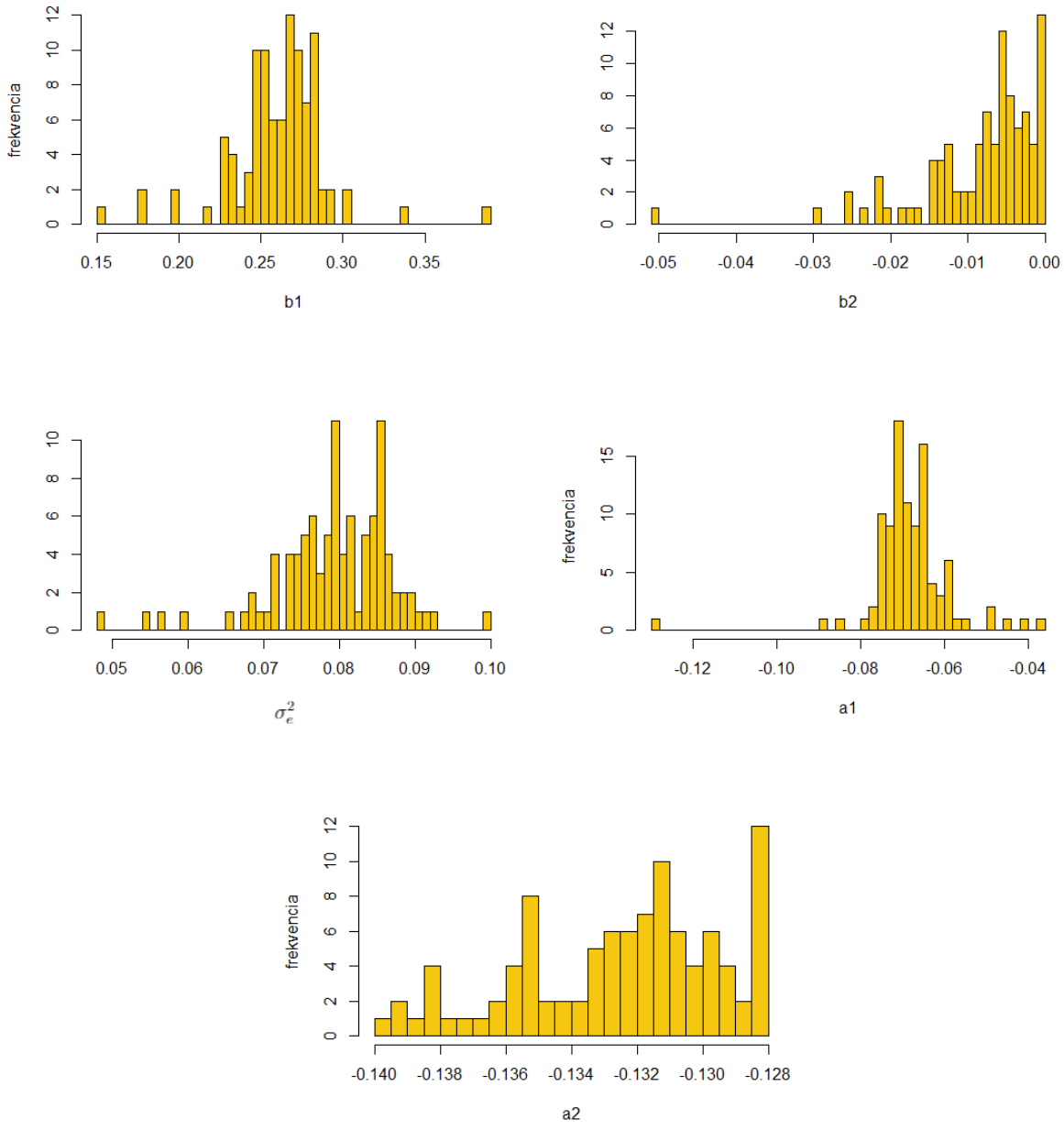
Tabuľka 8: Základné štatistiky pre výsledné hodnoty optimalizačných funkcií



Obr. 29: Pareto riešenia pre modelovanie európskych a domácich úrokových mier

Pozrieme sa na histogramy získaných parametrov pre Pareto optimálne riešenia (Obr.30). Parametre b_1, b_2 a σ_e^2 získavame ako riešenie optimalizačného problému. Následne parameter σ_d^2 bol zvolený ako štandardná odchýlka domácich úrokových mier s 3 mesačnou maturitou a parameter a_1 dopočítaný pomocou (32). Parameter a_2 sme získali optimalizačou funkciou *optim*, kde sme minimalizovali (31) vzhľadom na domáce úrokové mery.

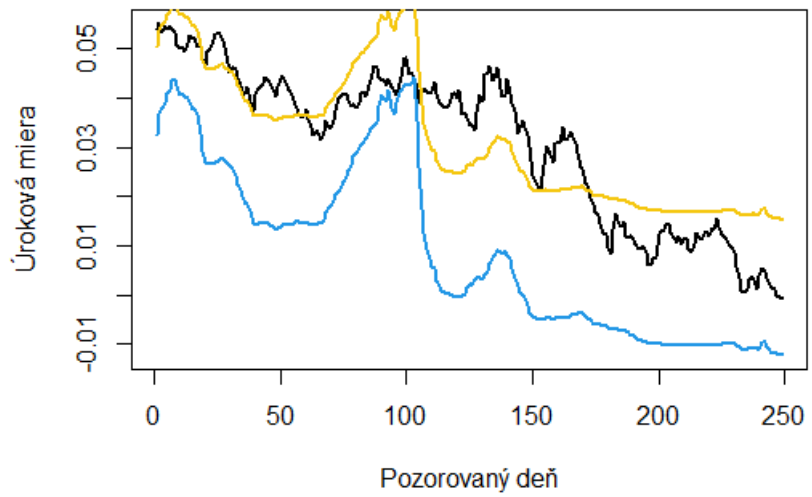
V danej optimalizácii sme nastavili ohraničenie pre parameter a_2 na interval [-5,-0.01]. Zápornosť premennej vychádza z charakteru samotnej premennej (podobne ako pri parametri b_2).



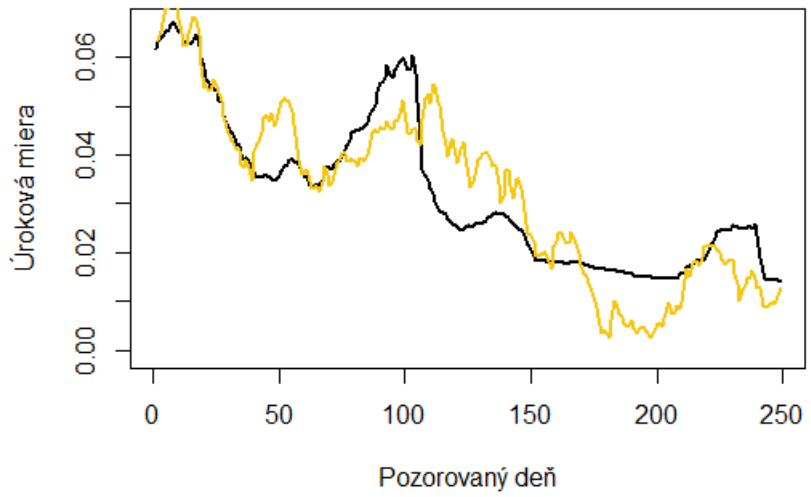
Obr. 30: Histogramy parametrov pre európske a české úrokové miery s 10 ročnou maturitou

Na Obr.31 a Obr.32 znázorňujeme, ako vyzerajú modelované úrokové miery vzhľadom na pozorované riešenia. Čierna farba prislúcha získaným dátam úrokových mier, modrá farba prislúcha modelovaným úrokovým mieram pre získané Pareto optimálne riešenie,

ktoré sme označili za outliera a následne žltá krivka predstavuje modelované úrokové miery pre náhodne zvolené iné Pareto optiálne riešenie (jedno z tých, ktoré sme ponechali ako efektívne riešenie). Modelované úrokové miery prislúchajúce outlierovi sa výraznejšie líšia iba pre európske úrokové miery. Domáce ostávajú takmer nezmenené. Túto skutočnosť môžeme vyčítať aj z hodnôt účelových funkcií (hodnota účelovej funkcie prislúchajúca domácej úrokovej miere má omnoho nižšiu disperziu ako hodnota účelovej funkcie prislúchajúca európskej úrokovej miere). Krivky modelovaných úrokových mier pre ostatné parametre zodpovedajúce Pareto optimálnym riešeniam sú takmer identické.



Obr. 31: Priebeh pozorovaných a modelovaných európskych úrokových mier s 10 ročnou maturitou



Obr. 32: Priebeh pozorovaných a modelovaných českých úrokových mier s 10 ročnou maturitou

Záver

Hlavným cieľom práce bolo využitím dvojkriteriálneho grey wolf algoritmu čo najlepšie odhadnúť parametre parametrických modelov, či parametre jednotlivých rozdelení dát.

V prvej časti práce sme podrobnejšie popísali myšlienku fungovania grey wolf genetického algoritmu a zároveň ozrejmili čitateľovi, ako chápať nájdené Pareto optimálne riešenia.

Spočiatku sme sa zamerali na nájdenie optimálnych parametrov pre pravdepodobnostné rozdelenia. Ako prvé sme nakalibrovali Burrovo rozdelenie a $LogPH$ rozdelenie, aby čo najlepšie popísalo dáta popisujúce škodové udalosti týkajúce sa požiarov v Dánsku. Pre Burrovo rozdelenie sme našli 6 Pareto optimálnych riešení a pre $LogPH$ 13 Pareto optimálnych riešení. Kým pre Burrovo rozdelnie hodnoty Komogorovej - Smirnovovej testovej štatistiky sa pohybovali v rozmedzí 0.0155 až 0.0175, v prípade $LogPH$ rozdelenia interval Komogorovej - Smirnovovej testovej štatistiky je ohraničený hodnotami 0.013 a 0.0135. Interval Cramér - von Missesovej testovej štatistiky sa zúžil z intervalu [0.106, 0.118] na interval [0.0807, 0.0817]. Vo vzájomnom porovnaní môžeme vzhľadom na dosiahnuté hodnoty testových štatistík konštatovať, že $LogPH$ rozdelenie popísuje dáta lepšie.

Druhá časť pravdepodobnostných rozdelení bola venovaná trhovému indexu $S\&P$ 500. Algoritmom sme sa pokúsili nakalibrovať Normálne a Variance gamma rozdelenie. V prípade Normálneho rozdelenia sme získali 446 Pareto optimálnych riešení, pre ktoré hustota rozdelenia omnoho lepšie kopíruje histogram dát, ako keby sme za strednú hodnotu a disperziu vzali priemer a varianciu dát. Trhový index omnoho lepšie popisovalo Variance gama rozdelenie, pre ktoré sa nám algoritmom podarilo získať 42 Pareto optimálnych riešení. Riešenia sa nezhodujú v jednom parametri γ. Daný parameter sa líšil minimálne a aj po grafickom vykreslení sme videli, že získané riešenia môžeme považovať za jedno. Aj zo získaných hodnôt testových štatistík, ktoré popisujú podobnosť kumulatívnej a kalibrovanej distribučnej funkcie, vidíme, že Variance gama rozdelenie omnoho lepšie fituje pozorované dáta. Kým minimálna hodnota K-S testovej štatistiky pre Normálne rozdelenie dosahovala hodnotu okolo 0.037, pre Variance gama rozdelenie to bolo okolo 0.013. Hodnota Cramér - von Missesovej testovej štatistiky sa pre Variance Gama rozdelenie znížila takmer desaťkrát.

V tretej časti práce sme hľadali optimálne váhy akcií pre portfólio, ktoré pozostávalo

zo štyroch najvýnosnejších akcií trhového indexu $S\&P$ 500. Váhy sme hľadali pre Markowitzov a Mean-Gini model. Výsledné riešenia Markowitzovho modelu sme porovnali s efektívnou hranicou portfólia, ktorú sme získali kvadratickou optimalizáciou rizika pre fixne stanovený výnos. Výsledné váhy v oboch modeloch vo väčšine prípadov značili o neinvestovaní do akcií Applu a Amazonu. Najväčšia časť investovaného kapitálu sa rozložila medzi akcie Invidii a Microsoftu.

V poslednej časti práce sme využitím viacfaktorového Vašíčkoveho modelu nakalibrovali domáce dlhodobé úrokové miery s maturitou 10 rokov (v našom prípade úrokové miery Českej republiky). Predpokladali sme, že tieto úrokové miery sú v rovnovážnom stave priťahované k európskym dlhodobým úrokovým mieram s maturitou 10 rokov, ktoré sme nakalibrovali jednofaktorovým Vašíčkovým modelom. Získanou kalibráciou sme dostali zopár outlierov, ktorých sme vyhodnotili ako nefektívne riešenia a zanedbali sme ich (kvôli vysokým hodnotám testovej štatistiky, ktorá popisovala rozdiely fitovaných a pozorovaných európskych úrokových mier). Pre zjednodušenie optimalizácie, algoritmom sme odhadli len parametre pre európske úrokové miery, následne parameter a_1 a pre domáce úrokové miery sme dopočítali a parameter a_2 získali ako výsledok jednorozmernej optimalizačnej funkcie *optim*.

Na záver konštatujeme, že sa nám daným dvojkriterialnym algoritmom podarilo nájsť veľmi dobré riešenia hľadaných parametrov. Získané výsledky dávali zmysel a zároveň dobre popisovali kalibrované dáta. V prípadoch, kedy sa v článkoch vyskytovali už nejaké optimálne parametre pre dané rozdelenie a dáta (ako tomu bolo v prípade dát o danských požiaroch), tak sme ukázali, že pre zvolené optimalizačné kritéria naše výsledky dávali minimálne rovnako dobré hodnoty testových štatistík ako získané výsledky v článku.

Je potrebné spomenúť, že výsledné parametre sú v značnej miere závislé od kvalitného počiatočného odhadu a primerane veľkom počte populácie. Pre niektoré odhady parametrov (napr. pre variance gamma rozdelenie) boli veľmi senzitívne na väčšiu zmenu intervalu prípustných riešení.

Ako nadstavbu práce by sme si vedeli predstaviť bližšie preskúmanie získaných Pareto optimálnych riešení. Veľakrát sa riešenia vzájomne líšili len o malé hodnoty. Bolo by užitočné spraviť hlbšiu analýzu a preskúmať, ako si zo všetkých riešení vybrať robustné riešenie. Viac sa takýmto robustným riešením venovali v [13].

Zoznam použitej literatúry

[1] Ahmad,Z., Mahmoudi, E., Hamedani, G. G., Kharazmi, O.: *New methods to define heavy-tailed distributions with applications to insurance data*, Journal of Taibah University for Science, Vol. 14 (2020), pp. 359-382

[2] Ahn, S., Kim, J.H., Ramaswami, V.: *A new class of models for heavy tailed distributions in finance and insurance risk*, Insurance: Mathematics and Economics, Vol. 51 (2012), pp. 43-52

[3] Al-Aboody, N.A., Al-Raweshidy, H.S.: *Grey wolf optimization-based energy-efficient routing protocol for heterogeneous wireless sensor networks*, in 2016 4th international symposium on computational and business intelligence, IEEE, 2016, pp. 101-107

[4] Behr, A., Pötter, U.: *Alternatives to the normal model of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models*, Annals of Finance, Vol. 5 (2009), pp. 49-68

[5] Bhensdadia, V., Tejani, G.: (2016). *Grey wolf optimizer (GWO) algorithm for minimum weight planer frame design subjected to AISC-LRFD*, in Proceedings of International Conference on ICT for Sustainable Development, Springer, Singapore, 2016, pp. 143-151

[6] Bučková, Z., Girová, Z., Stehlíková, B.: *Estimating the domestic short rate in a convergence model of interest rates*, Tatra Mountains Mathematical Publications, Vol. 75 (2020), 33-48

[7] Chandra, M., Agrawal, A., Kishor, A., Niyogi, R.: *Web service selection with global constraints using modified gray wolf optimizer*, in 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2016, pp. 1989-1994

[8] Chaman-Motlagh, A.: Superdefect photonic crystal filter optimization using grey wolf optimizer, IEEE Photonics Technology Letters, Vol. 27 (2015), pp. 2355-2358

[9] Choi, S.Y., Yoon, J.H.: *Modeling and risk analysis using parametric distributions with an application in equity-linked securities*, Mathematical Problems in Engineering, 2020

[10] Chutima, P., Pinkoompee, P. : *Multi-objective sequencing problems of mixed-model assembly systems using memetic algorithms*, Science Asia, Vol. 35 (2009), pp. 295-305

[11] Corzo Santamaria, T., Schwartz, E. S.: *Convergence within the EU: Evidence from interest rates*, Economic Notes, Vol. 29 (2000), pp. 243-266

[12] Dánske dáta o požiaroch http://www.ma.hw.ac.uk/∼ mcneil/data.html

[13] Deb, K., Gupta, H.: *Searching for robust Pareto-optimal solutions in multi-objective optimization*, In Evolutionary Multi-Criterion Optimization, Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005, Proceedings 3, Springer Berlin Heidelberg, 2005, pp. 150-164

[14] Dianová, P.: *Aplikácia heuristických optimalizačných algoritmov na kalibráciu modelov úrokových mier*, diplomová práca, FMFI UK, Bratislava, 2021, dostupné na internete (1. 1. 2023) www.iam.fmph.uniba.sk/studium/efm/diplomovky/2021/dianova/diplomovka.pdf

[15] Doumpos, M., Zopounidis, C.: *Multi-objective optimization models in finance and investments*, J Glob Optim, Vol. 76 (2020), pp. 243–244

[16] El-Fergany, A.A., Hasanien, H.M.: *Single and multi-objective optimal power flow using grey wolf optimizer and differential evolution algorithms*, Electric Power Components and Systems, Vol. 43 (2015), 1548-1559.

[17] Evans, D. L., Drew, J. H., Leemis, L. M.: *The distribution of the Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling test statistics for exponential populations with estimated parameters*, Communications in Statistics—Simulation and Computation®, Vol. 37 (2008), pp. 1396-1421

[18] Gould, W., Pitblado, J., Sribney, W.: *Maximum likelihood estimation with Stata*, Stata press, 2006

[19] Grey wolf optimalization, dostupné na internete (1. 1. 2023) https://seyedalimirjalili.com/gwo

[20] Hadidian-Moghaddam, M.J., Arabi-Nowdeh, S., Bigdeli, M.: *Optimal sizing of a stand-alone hybrid photovoltaic/wind system using new grey wolf optimizer considering reliability*, Journal of Renewable and Sustainable Energy, Vol. 8 (2016), 035903

[21] Jašurková, T., Stehlíková, B.: *Calibration of the Vasicek model of interest rates using bicriteria optimization*, Proceedings of ALGORITMY, 2020, pp. 211-220

[22] Ji, R., Lejeune, M. A., Prasad, S. Y.: *Interactive portfolio optimization using Mean-Gini criteria*. Financial Decision Aid Using Multiple Criteria: Recent Models and Applications, 2018, pp. 49-91

[23] Keller, A. A.: *Multi-objective optimization in theory and practice II: metaheuristic algorithms*, Bentham Science Publishers, 2019

[24] Krink, T., Paterlini, S.: *Multiobjective optimization using differential evolution for real-world portfolio optimization*, Computational Management Science, Vol. 8 (2011), pp. 157-179

[25] Melicherčík, I., Olšarová, L., Úradníček, V.: *Kapitoly z finančnej matematiky 1*, Bratia Sabovci, 2005

[26] Mirjalili, S., Mirjalili, S. M., Lewis, A.: *Grey Wolf Optimizer*, Advances in Engineering Software, Vol. 69 (2014), pp. 46-61

[27] Mirjalili, S., Saremi, S., Mirjalili, S. M., Coelho, L. D. S.: *Multi-objective grey wolf optimizer: a novel algorithm for multi-criterion optimization*, Expert Systems With Applications, Vol. 47 (2016), pp. 106-119

[28] Rangaiah, G. P. (Ed.): *Multi-objective optimization: techniques and applications in chemical engineering*, world scientific, Vol. 5 (2016)

[29] Szűcs, G.: *Matematické modelovanie v poisťovníctve*, elektronické skripta

[30] Úrokové miery - dáta dostupné (1.4.2023) data.oecd.org/interest/short-term-interest-rates

[31] Van Griensven, A., Meixner, T: *A global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models*, Journal of Hydroinformatics, Vol. 9 (2007), pp.277-291

[32] Výber akcií $S\&P$ 500 - dáta dostupné (1.4.2023) slickcharts.com/sp500)

[33] Quantmod - balík pre získanie dát pre akcie $S\&P$ 500 https://cran.r-project.org/web/packages/quantmod/index.html

[34] Seneta, E.: *Fitting the variance-gamma model to financial data*, Journal of Applied Probability, Vol. 41 (2004), pp. 177-187

[35] Shieh, C.S., Wang, H.Y., Dao, T.K.: *Enhanced diversity herds grey wolf optimizer for optimal area coverage in wireless sensor networks*, in International conference on genetic and evolutionary computing, Springer, Cham, 2016, pp. 174-182

Príloha A

```
############Danske data#############
########Burrovo rozdelenie##########

set.seed(123456)

fobj = function(parametre){
  #Parametre:
  #shape1 = parametre[1],
  #shape2 = parametre[2],
  #scale = parametre[3]

  #vycislenie distr. fcii pre data
  VDF <- pburr(danish_data, shape1 = parametre[1],
  shape2 = parametre[2], scale = parametre[3])

  #vycislenie empirickej fcii pre data
  ECDF <- ecdf(danish_data)
  EDF <- ECDF(danish_data)

  # ll - Hodnota druhej optimalizacnej fcii -
  # K-S testova statistika (14)
  vz1 <- abs(VDF - EDF)
  vz2 <- abs(VDF[2:N] - EDF[1:(N-1)])
  ll=  max(vz1, vz2)

  # TT - Hodnota prvej optimalizacnej fcii -
  # Cramer-von Mises testova statistika (15)

  TT = 1/(12*N) + sum((((2*c(1:N) - 1)/(2*N)) - VDF)^2)
```

```
  cbind(ll,TT)
}

xrange = matrix(1,nVar,2)
# Ohranicenie parametrov
for (i in 1:nVar) {
  #shape1
  xrange[1,1] = 0.01
  xrange[1,2] = 5

  #shape2
  xrange[2,1] = 0.01
  xrange[2,2] = 5

  #scale
  xrange[3,1] = 0.01
  xrange[3,2] = 5
}

lb = xrange[,1]
ub = xrange[,2]

#definovanie parametrov algoritmu
VarSize  = c(1,nVar)
GreyWolves_num=500
MaxIt=10
Archive_size=1500
alpha=0.1
nGrid=10
beta=4
gamma=2
```

Príloha B

```
############Danske data#############
#########LogPH rozdelenie##########

set.seed(123456)
#definovanie optimalizacnej funkcie
fobj = function(parametre){
  #parametre:
  #Matica T - T,
  #vektor alpha - aalpha,
  #jednotkovy vektor - one
  T = matrix(c(parametre[1], parametre[2],
  parametre[3], parametre[4]), nrow = 2)
  aalpha = c(parametre[5], 1 - parametre[5])
  one = rep(1, dim(T)[2])

  #podmienka pripusnosti parametrov
  if (Re(eigen(T)$values[1])<0 & Re(eigen(T)$values[2])<0){

    # prepis distr. fcia rozdelenia (17)
    disfun = function(y){
      DD = aalpha%*%expm(T*log(y))%*%one
      disfun = 1 - DD[1,1]
    }

    #vycislenie distr. fcii pre data
    VDF <- c()
    for (i in 1:N) {
      VDF[i] <- disfun(danish_data[i])
    }
```

```
    #vycislenie empirickej fcii pre data
    ECDF <- ecdf(danish_data)
    EDF <- ECDF(danish_data)

    # ll - Hodnota prvej optimalizacnej fcii -
    # K-S testova statistika (14)
    vz1 <- abs(VDF - EDF)
    vz2 <- abs(VDF[2:N] - EDF[1:(N-1)])
    ll=  max(vz1, vz2)

    # TT - Hodnota druhej optimalizacnej fcii -
    # Cramer-von Mises testova statistika (15)
    TT = 1/(12*N) + sum((((2*c(1:N) - 1)/(2*N)) - VDF)^2)

  }else{
    ll = Inf
    TT = Inf

  }
  cbind(ll,TT)

}

xrange = matrix(1,nVar,2)

# Ohranicenie parametrov
for (i in 1:nVar) {
  #matica T[1,1]
  xrange[1,1] = -5
  xrange[1,2] = -0.01
```

```
  #matica T[2,1]
  xrange[2,1] = 0.01
  xrange[2,2] = 5

  #matica T[1,2]
  xrange[3,1] = 0.01
  xrange[3,2] = 5

  #matica T[2,2]
  xrange[4,1] = -5
  xrange[4,2] = -0.01

  #alpha[1]
  xrange[5,1] = 0
  xrange[5,2] = 1

}

lb = xrange[,1]
ub = xrange[,2]

#definovanie parametrov algoritmu
VarSize  = c(1,nVar)
GreyWolves_num=10000
MaxIt=10
Archive_size=500
alpha=0.1
nGrid=10
beta=4
gamma=2
```

Príloha C

```
############Data S&P 500############
#########Normalne rozdelenie##########

set.seed(123456)

fobj = function(parametre){
  #Parametre:
  #mean = parametre[1],
  #sd = parametre[2]

  #vycislenie distr. fcii pre data
  VDF <- pnorm(SP, mean = parametre[1], sd = parametre[2])

  #vycislenie empirickej fcii pre data
  ECDF <- ecdf(SP)
  EDF <- ECDF(SP)

  # ll - Hodnota prvej optimalizacnej fcii -
  # K-S testova statistika (14)
  vz1 <- abs(VDF - EDF)
  vz2 <- abs(VDF[2:N] - EDF[1:(N-1)])
  ll=  max(vz1, vz2)

  # TT - Hodnota druhej optimalizacnej fcii -
  # Cramer-von Mises testova statistika (15)

  TT = 1/(12*N) + sum((((2*c(1:N) - 1)/(2*N)) - VDF)^2)

  cbind(ll,TT)
```

```
}


xrange = matrix(1,nVar,2)

# Ohranicenie parametrov
for (i in 1:nVar) {
  #mu
  xrange[1,1] = min(SP)
  xrange[1,2] = max(SP)

  #sigma
  xrange[2,1] = 0
  xrange[2,2] = 1
}

lb = xrange[,1]
ub = xrange[,2]

#definovanie parametrov algoritmu
VarSize  = c(1,nVar)
GreyWolves_num=500
MaxIt=10
Archive_size=50000
alpha=0.1
nGrid=10
beta=4
gamma=2
```

Príloha D

```
############Data S&P 500############
#####Variance Gamma rozdelenie######

set.seed(35421)

fobj = function(parametre){
  #Parametre:
  #vgC = parametre[1]
  #sigma =  parametre[2]
  #theta = parametre[3]
  #nu = parametre[4]

  #vycislenie distr. fcii pre data
  VDF <- pvg(SP, vgC = parametre[1],sigma =  parametre[2],
  theta = parametre[3], nu = parametre[4])

  #vycislenie empirickej fcii pre data
  ECDF <- ecdf(SP)
  EDF <- ECDF(SP)

  # ll - Hodnota prvej optimalizacnej fcii -
  # K-S testova statistika (14)
  vz1 <- abs(VDF - EDF)
  vz2 <- abs(VDF[2:N] - EDF[1:(N-1)])
  ll=  max(vz1, vz2)

  # TT - Hodnota druhej optimalizacnej fcii -
  # Cramer-von Mises testova statistika (15)

  TT = 1/(12*N) + sum((((2*c(1:N) - 1)/(2*N)) - VDF)^2)
```

```
  cbind(ll,TT)
}

xrange = matrix(1,nVar,2)

# Ohranicenie parametrov
for (i in 1:nVar) {
  #vgC
  xrange[1,1] = 0.0002585*2
  xrange[1,2] = 0.0002585/2

  #sigma
  xrange[2,1] = 0.008029321*2
  xrange[2,2] = 0.008029321/2

  #theta
  xrange[3,1] = -0.000151/2
  xrange[3,2] = -0.000151*2

  #nu
  xrange[4,1] = 0.4220*2
  xrange[4,2] = 0.4220/2
}

lb = xrange[,2]
ub = xrange[,1]

#definovanie parametrov algoritmu
VarSize  = c(1,nVar )
GreyWolves_num=500
```

MaxIt=10

Archive_size=500

alpha=0.1

nGrid=10

beta=4

gamma=2

Príloha E

```
#####Optimalizacia portfolia#####
#########Markowitz model##########

#funkcia pre denne vynosy zvolenych akcii
get_daily_return <- function(assets) {

  # Stiahnutie historickych cien
  prices <- lapply(assets, function(x) {
    getSymbols(x, from = "2015-01-01",
               to = "2023-01-01", auto.assign = TRUE)
    Ad(get(x))
  })

  prices <- na.exclude(prices)

  # Prekonvertovanie na xts objekt
  prices_xts <- do.call(merge, lapply(prices, as.xts))

  # Vypocet dennych vynosov
  returns_daily <- lapply(prices_xts, Return.calculate,
                          method = "log")
  names(returns_daily) <- assets
  returns_daily <- do.call(cbind, returns_daily)
  returns_daily <-na.omit(returns_daily)

  return(returns_daily)
}

#funkcia pre priemerne rocne vynosy zvolenych akcii
get_yield_return <- function(assets) {
```

```
  # Stiahnutie historickych cien
  prices <- lapply(assets, function(x) {
    getSymbols(x, from = "2015-01-01",
               to = "2023-01-01", auto.assign = TRUE)
    Ad(get(x))
  })
  prices <- na.exclude(prices)

  # Prekonvertovanie na xts objekt
  prices_xts <- do.call(merge, lapply(prices, as.xts))

  # Vypocet dennych vynosov
  returns_daily <- lapply(prices_xts, Return.calculate,
                          method = "log")
  returns_daily <-na.omit(returns_daily)
  returns_year <- lapply(returns_daily, Return.annualized,
                         scale = 255)
  names(returns_daily) <- assets
  returns_year <- do.call(cbind, returns_year)

  return(returns_year)
}


#vyber akcii
assets <- c("AAPL","MSFT", "AMZN","NVDA")

#Rocne vynosy akcii
returns_yield <- get_yield_return(assets)
```

```
#Denne vynosy akcii
returns_daily <- get_daily_return(assets)

returns_yield <- as.vector(returns_yield)
returns_daily <- as.matrix(returns_daily)
sample_data <- data.frame( var1 = returns_daily[,1],
                           var2 = returns_daily[,2],
                           var3 = returns_daily[,3],
                           var4 = returns_daily[,4])

#Kovariancna matica vynosov
COV_MAT <- cov(sample_data)
#vektor priemernych rocnych vynosov
YIELD_RET <- returns_yield

set.seed(123456)

fobj = function(parametre){

  # vahy pre 1. 2. a 3. akciu: parametre[1], parametre[2],
  # parametre[3]
  # vaha pre 4. akciu: dopocitana (1 - vypocitane vahy)
  pom <- c(parametre[1], parametre[2], parametre[3])
  pom <- sum(pom)
  vahy <- c(parametre[1], parametre[2], parametre[3], 1 - pom)

  if (pom < 1){
    # 1. Optimalizacna funkcia - minimalizacia rizika portfolia (23)
    TT = vahy%*%COV_MAT%*%vahy

    # 2. Optimalizacna funkcia - maximalizacia vynosu portfolia (24)
```

```
    ll = -(vahy%*%YIELD_RET)
  }else{
    ll = Inf
   TT = Inf
  }

  cbind(ll, TT)
}


xrange = matrix(1,nVar,2)

# Ohranicenie parametrov
for (i in 1:nVar) {
  #vaha pre 1.akciu
  xrange[1,1] = 0
  xrange[1,2] = 1

  #vaha pre 2.akciu
  xrange[2,1] = 0
  xrange[2,2] = 1

  #vaha pre 3.akciu
  xrange[3,1] = 0
  xrange[3,2] = 1

}

lb = xrange[,1]
ub = xrange[,2]
```

```
#definovanie parametrov algoritmu
VarSize  = c(1,nVar)
GreyWolves_num=500
MaxIt=10
Archive_size=50000
alpha=0.1
nGrid=10
beta=4
gamma=2
```

Príloha F

```
#####Optimalizacia portfolia#####
#########Mean-gini model##########

#funkcia pre denne vynosy zvolenych akcii
get_daily_return <- function(assets) {

  # Stiahnutie historickych cien
  prices <- lapply(assets, function(x) {
    getSymbols(x, from = "2015-01-01", to = "2023-01-01",
               auto.assign = TRUE)
    Ad(get(x))
  })

  prices <- na.exclude(prices)

  # Prekonvertovanie na xts objekt
  prices_xts <- do.call(merge, lapply(prices, as.xts))

  # Vypocet dennych vynosov
  returns_daily <- lapply(prices_xts, Return.calculate,
                          method = "log")
  names(returns_daily) <- assets
  returns_daily <- do.call(cbind, returns_daily)
  returns_daily <-na.omit(returns_daily)

  return(returns_daily)
}

#funkcia pre priemerne rocne vynosy zvolenych akcii
get_yield_return <- function(assets) {
```

```
  # Stiahnutie historickych cien
  prices <- lapply(assets, function(x) {
    getSymbols(x, from = "2015-01-01", to = "2023-01-01",
               auto.assign = TRUE)
    Ad(get(x))
  })
  prices <- na.exclude(prices)

  # Prekonvertovanie na xts objekt
  prices_xts <- do.call(merge, lapply(prices, as.xts))

  # Vypocet dennych vynosov
  returns_daily <- lapply(prices_xts, Return.calculate,
                          method = "log")
  returns_daily <-na.omit(returns_daily)
  returns_year <- lapply(returns_daily, Return.annualized,
                         scale = 255)
  names(returns_daily) <- assets
  returns_year <- do.call(cbind, returns_year)

  return(returns_year)
}

#funkcia pre priemerne tyzdenne vynosy zvolenych akcii
get_weekly_return <- function(assets) {

  # Stiahnutie historickych cien
  prices <- lapply(assets, function(x) {
    getSymbols(x, from = "2015-01-01", to = "2023-01-01",
               auto.assign = TRUE)
```

```
    Ad(get(x))
  })
  prices <- na.exclude(prices)

  # Prekonvertovanie na xts objekt
  prices_xts <- do.call(merge, lapply(prices, as.xts))

  # Vypocet tyzdennych vynosov
  returns_weekly <- lapply(prices_xts, weeklyReturn,
                           method = "log")
  returns_weekly <-na.omit(returns_weekly)
  names(returns_weekly) <- assets
  returns_weekly <- do.call(cbind, returns_weekly)

  return(returns_weekly)
}

#vyber akcii
assets <- c("AAPL","MSFT", "AMZN","NVDA")

#rocne, tyzdenne a denne vynosy
returns_yield <- get_yield_return(assets)
returns_weekly <- get_weekly_return(assets)
returns_daily <- get_daily_return(assets)

returns_yield <- as.vector(returns_yield)
returns_weekly <-as.matrix(returns_weekly)
returns_daily <-as.matrix(returns_daily)

YIELD_RET <- returns_yield
M <- returns_weekly
```

```
M <- returns_weekly[2:dim(M)[1],]
M <- t(M)


set.seed(123456)

fobj = function(parametre){

  # vahy pre 1. 2. a 3. akciu: parametre[1], parametre[2],
  # parametre[3]
  # vaha pre 4. akciu: dopocitana (1 - vypocitane vahy)
  pom <- c(parametre[1], parametre[2], parametre[3])
  pom <- sum(pom)
  vahy <- c(parametre[1], parametre[2], parametre[3], 1 - pom)

  if (pom < 1){

    #1. optimalizacna funkcia - min. Mean-Gini kriterium (25)
    TT <-0
    for (t in 1:(dim(M)[2]-1)) {
      for (tt in (t+1):dim(M)[2]) {
        for (j in 1:dim(M)[1]) {
          TT1 <- abs(vahy[j]*M[j,t]-vahy[j]*M[j,tt])
          TT <- TT+TT1
        }
      }
    }

    #1. optimalizacna funkcia - max. vynos portfolia
    ll = -(vahy%*%YIELD_RET)
  }else{
```

```
    ll = Inf
    TT = Inf
  }

  cbind(ll, TT)
}


xrange = matrix(1,nVar,2)

# Ohranicenie parametrov
for (i in 1:nVar) {
  #vaha pre 1.akciu
  xrange[1,1] = 0
  xrange[1,2] = 1

  #vaha pre 2.akciu
  xrange[2,1] = 0
  xrange[2,2] = 1

  #vaha pre 3.akciu
  xrange[3,1] = 0
  xrange[3,2] = 1

}

lb = xrange[,1]
ub = xrange[,2]

#definovanie parametrov algoritmu
VarSize  = c(1,nVar)
```

GreyWolves_num=500

MaxIt=10

Archive_size=50000

alpha=0.1

nGrid=10

beta=4

gamma=2

Príloha G

```
##########Urokove miery###########
#########Vasickov model##########

#data:
#SRD - short rate interest rate domestic
#LRD - long rate interest rate domestic
#SRE - short rate interest rate euro
#LRE - long rate interest rate euro

#b1 - parametre[1], b2 - parametre[2], sigmae^2 - parametre[3]
set.seed(12345)

fobj = function(parametre){

  #fcia pre vypocet europskej urokovej miery
  # podla Vasickovho modelu
  modelLREFun = function(Tau, b1, b2, sigmae2){
    #fcia G(Tau) (31)
    GG = function(Tau, b2){
      GG = (1 - exp(b2*Tau))/(-b2)
    }

    #fcia G(Tau) (32)
    FF = function(Tau, b1, b2, sigmae2){
      FF = ((GG(Tau, b2) - Tau)*((-b2)*b1 - sigmae2/2))/b2^2
           - (sigmae2*GG(Tau, b2)^2)/(-4*b2)
    }

    Ffun <- FF(Tau, b1, b2, sigmae2)
    Gfun <- GG(Tau, b2)
```

```
  #vypocet europskej urokovej miery(30)
  modelLRE = (-1/Tau)*(Ffun - SRE*Gfun)
}

#funkcia pre vypocet domacej urokovej miery
modelLRDFun = function(Tau, b1, b2, sigmae2, a2){

  #funkcia D(Tau)
  DD = function(Tau, a2){
    DD = (- 1 + exp(a2*Tau))/a2
  }

  #funkcia E(Tau)
  EE = function(Tau, a2, b2){
    EE = (-1/b2)*((a2*(1 - exp(b2*Tau))
         - b2*(1 - exp(a2*Tau)))/(a2 - b2))
  }

  #funkcie F1, F2 a F3 ako vycislene integraly
  FF1 = (a2*Tau - exp(a2*Tau) + 1)/a2^2
  FF2 = (2*a2*Tau - 4*exp(a2*Tau) + exp(2*a2*Tau) + 3)/(4*a2^3)
  FF3a =  -(b1 *(-(b2 *(exp(a2*Tau) - 1))/a2 +
           (a2 *(b2 *(-Tau) + exp(b2* Tau) - 1))/b2 +
           b2*Tau))/(b2 *(a2 - b2))
  FF3b = (sigmae2 *(a2^4 *(2*b2 *Tau - 4*exp(b2 *Tau) +
          exp(2*b2*Tau) + 3) - a2^3* b2 *(2* b2 *Tau -
          exp(2* b2 *Tau) + 1) - 2 *a2^2 *b2^2 *
          (2 *(exp(a2 *Tau) - 1) *(exp(b2* Tau) - 1) +
          b2 *Tau) + a2*b2^3 *(exp(2 *a2 *Tau) + 2* b2 *Tau - 1) +
          b2^4* (-4* exp(a2* Tau) + exp(2 *a2 *Tau) +
```

```
            3)))/(4 *a2 *b2^3 *(a2 - b2)^2 *(a2 + b2))
  FF3 = FF3a + FF3b

  y = -Tau*LRD - FF3 + DD(Tau, a2)*SRD + EE(Tau, a2, b2)*SRE
  sigmad2 = var(SRD)
  #zo vztahu (36), vypocet parametra a1
  a1 = sum(y)/(length(y)*FF1) - sigmad2*FF2/FF1

  #model pre domace urokove miery
  modelLRD = (a1*FF1 + sigmad2*FF2 + FF3 - DD(Tau, a2)*SRD -
              EE(Tau, a2, b2)*SRE)*(-1/Tau)
  return(list(modelLRD = modelLRD, a1 = a1))
}

b1 = parametre[1]
b2 = parametre[2]
sigmae2 = parametre[3]

#vasicek - model jednofaktorovy
#1. ucelova fcia - vztah (35) pre domace urokove miery
LRE_model = modelLREFun(Tau, b1, b2, sigmae2)
pom1 <- LRE_model - LRE
ll = sum(pom1^2)

#model domestic rates
modelSRDPom = function(a2){

  #vasicek pre domace - dvojfaktorovy
  #2. ucelova fcia - vztah (35) pre europske urokove miery
  modelLRD = modelLRDFun(Tau, b1, b2, sigmae2, a2)$modelLRD
  pom2 <- (modelLRD - LRD)
```

```
    TT = sum(pom2^2)

  }
  #vypocet parametra a2
  op <- optim(0.001, modelSRDPom, method = "L-BFGS-B",
              lower = -5, upper = -0.01)
  a2 <- op$par
  TT <- op$value

  cbind(ll,TT)

}

xrange = matrix(1,nVar,2)

# Ohranicenie parametrov
for (i in 1:nVar) {
  #b1
  xrange[1,1] = -0.5
  xrange[1,2] = 0.5

  #b2
  xrange[2,1] = -0.1
  xrange[2,2] = 0

  #sigmae2
  xrange[3,1] = 0
  xrange[3,2] = 0.1

}
```

```
lb = xrange[,1]
ub = xrange[,2]

#definovanie parametrov algoritmu
VarSize  = c(1,nVar )
GreyWolves_num=1000
MaxIt=10
Archive_size=5000
alpha=0.1
nGrid=10
beta=4
gamma=2
```