Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Andrea Polakovičová

Úloha replikácie indexov pomocou mier rizika

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	Doc. RNDr. Martin Branda, Ph.D.
Studijní program:	Matematika
Studijní obor:	Finanční Matematika

Praha 2023

Touto formou by som sa rada poďakovala môjmu vedúcemu bakalárskej práce Doc. RNDr. Martinovi Brandovi, Ph.D. za jeho prístup, cenné rady a čas, ktorý mi venoval v dobe vypracovávania tejto práce.

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

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

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

Názov práce: Úloha replikácie indexov pomocou mier rizika

Autor: Andrea Polakovičová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: Doc. RNDr. Martin Branda, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V tejto práci sa budeme zaoberať jednotlivými metódami merania rizík známymi pod pojmami Value at Risk (VaR) a Conditional Value at Risk ($CVaR$). Ich vlastnosti a formulácie využijeme pri odvodení optimalizačnej úlohy lineárneho typu. Úloha lineárneho programovania bude pozostávať z minimalizácie účelovej funkcie predstavujúcej odchýlku medzi portfóliom a zvoleným indexom. Výpočet prevedieme na základe viacerých podmienok, kde práve jedna z nich bude využívať uvedené metódy merania rizika VaR a $CVaR$. Cieľom je na základe tohto programu vytvoriť portfólio, pomocou ktorého budeme replikovať index S&P 500. Celý výpočet prevedieme v programe Python na základe historických dát. Následne použijeme optimálne riešenie, ktoré software našiel a zostavíme podľa neho replikačné portfólio, ktorého vývoj budeme sledovať v nasledujúcich časových obdobiach. V závere práce budeme analyzovať a diskutovať jednotlivé výsledky pre rôzne vstupné parametre.

Kľúčové slová: Miery rizika, optimalizácia portfólia, replikácia indexu

Title: Index tracking problem using risk measures

Author: Andrea Polakovičová

Department: Department of probability and mathematical statistics

Supervisor: Doc. RNDr. Martin Branda, Ph.D., Department of probability and mathematical statistics

Abstract: In this thesis, we will introduce various methods for measuring risk known as Value at Risk (VaR) and Conditional Value at Risk ($CVaR$). We will use their properties and formulations in deriving a linear optimization problem. The linear programming problem will consist of minimizing the objective function representing the deviation between the portfolio and a chosen index. The calculation will be carried out based on multiple constraints, where one of them will use the aforementioned risk measures VaR and $CVaR$. The goal is to create a portfolio based on this program that replicates the S&P 500 index. We will perform the entire calculation using Python based on historical data. Subsequently, we will use the optimal solution found by the software and construct a replication portfolio that we will track in the following time periods. In conclusion, we will analyze and discuss the individual results for various input parameters.

Keywords: Risk measures,portfolio optimization,index tracking

Obsah

Úvod

V dnešnej dobe sme všetci tlačení k tomu, aby sme naše peniaze investovali do rôznych finančných aktív vzhľadom na vysokú mieru inflácie. Avšak, každá investícia prináša so sebou riziko, ktoré sa neustále mení v závislosti od komplexnosti finančných nástrojov, do ktorých peniaze investujeme.

V tejto práci si predstavíme, čo pojem riziko predstavuje a zoznámime sa s metódami známymi ako Value at Risk (VaR) a Conditional Value at Risk ($CVaR$), pomocou ktorých vieme dané riziko merať a riadiť. Zameriame sa na ich vlastnosti, spôsoby ich výpočtu a jednotlivé súvislosti, ktoré medzi sebou miery majú.

Na základe ich vlastností odvodíme funkciu, pomocou ktorej budeme mať možnosť vypočítať obe miery rizika zároveň. Preskúmame jednotlivé vlastnosti odvodenej funkcie, ktoré budú pre nás neskôr veľmi užitočné.

Cieľom tejto práce bude skonštruovať portfólio, ktoré bude replikovať zvolený finančný index, keďže indexy sú považované z dlhodobého hľadiska za výnosné a mali by pre nás predstavovať menšiu rizikovosť. Na tento účel využijeme znalosti optimalizácie a zostavíme lineárny program, ktorého jedna z podmienok bude práve obmedzenie na mieru rizika. K formulácií tohto obmedzenia využijeme našu odvodenú funkciu, ktorú budeme obmedzovať zhora, vzhľadom na to akú výšku rizika budeme ochotní tolerovať.

Súčasťou výstupu lineárneho programu budú aj optimálne hodnoty jednotlivých aktív, z ktorých bude portfólio pozostávať. Výpočty budeme vykonávať na základe historických dát pomocou programovacieho jazyka Python.

Následne sa budeme venovať pozorovaniu vývoja výsledného portfólia, ktoré bude skonštruované na základe optimálnych hodnôt a jeho správaniu vzhľadom na zvolený index v nasledujúcich časových obdobiach. Týmto spôsobom sa pokúsime lepšie porozumieť rizikovosti nášho investičného portfólia a nahliadneme na presnosť jednotlivých výsledkov.

Kapitola 1

Miery rizika a ich vlastnosti

V tejto kapitole si uvedieme základné pojmy používané v oblasti financií. Zavedieme ich matematické definície, vyjadrenia a objasníme ich interpretáciu. Uvedieme čo je, a v akom zmysle je v našom prípade chápané riziko. Ďalej sa dostaneme k niektorým užitočným vlastnostiam funkcií, ktoré budeme neskôr využívať.

1.1 Základné definície

Definícia 1 (Portfólio)**.** *(Dupačová (2002)) Nech $n \in N$ je počet aktív, ktoré máme k dispozícií. Potom povieme, že vektor $\boldsymbol{x} = (x_1, \dots, x_n)$ je portfólio ak $\sum_{i=1}^{n} x_i = 1$ a zároveň pre $\forall i$, kde $i = 1, \dots, n$ platí $x_i \geq 0$.*

Na základe definície 1, portfólio bude pre nás predstavovať vektor o dĺžke n, kde n zodpovedá počtu finančných aktív, napríklad akcií do ktorých chceme investovať. Ďalej si môžme predstaviť, že množstvo peňazí, ktoré budeme investovať je rovné hodnote 1. Potom práve $\sum_{i=1}^{n} x_i = 1$, nám hovorí o tom, že jednotlivé hodnoty $x_1, \dots, x_n$ reprezentujú časť našej počiatočnej investície, ktorá je vložená do konkrétneho aktíva, kde musí platiť že celkový súčet investícií do jednotlivých aktív zodpovedá presne hodnote našej počiatočnej investície. Keďže zároveň musí pre všetky i platiť $x_i \geq 0$, budeme uvažovať také portfólio, v ktorom nie sú povolené predaje nakrátko.

K zavedeniu následujúcej definície budeme uvažovať pravdepodobnostný priestor $(Y, \mathcal{A}, P)$, kde $\mathcal{A}$ je σ-algebra na Y a P je pravdepodobnostná miera na Y. Zároveň budeme uvažovať reálny náhodný vektor $\boldsymbol{y} : \{Y, \mathcal{A}\} \rightarrow \{\mathbb{R}^n, \mathcal{B}_0^n\}$, u ktorého budeme predpokladať nezávislosť na predom danom vektore $\boldsymbol{x}$, spĺňajúceho definíciu 1.

Definícia 2 (Strata)**.** *(Rockafellar a Uryasev (2002)) Nech $\boldsymbol{x}$ je predom daný vektor vybraný z určitej podmnožiny $X \in \mathbb{R}^n$, potom $z = f(\boldsymbol{x}, \boldsymbol{y})$ je náhodná strata vzhľadom k rozhodovaciemu vektoru $\boldsymbol{x}$ a náhodnému vektoru $\boldsymbol{y}$ a zároveň platí, že funkcia $f(\boldsymbol{x}, \boldsymbol{y})$ je spojitá v $\boldsymbol{x}$ a merateľná v $\boldsymbol{y}$.*

V definícií 2 budeme vektor $\boldsymbol{x}$ interpretovať ako vektor reprezentujúci konkrétne portfólio v zmysle definície 1 a množinu X, ako množinu, ktorá predstavuje všetky dostupné portfólia. Náhodný vektor $\boldsymbol{y}$ bude predstavovať určité neistoty, ktoré môžu hodnotu straty ovplyvniť ako napríklad náhodné ceny uvažovaných aktív. Všeobecne budeme uvažovať, že záporná hodnota straty z predstavuje zisk.

Definícia 3 (Distribučná funkcia straty). *(Rockafellar a Uryasev (2002)) Pre všetky* $\boldsymbol{x} \in X$*, označíme* $\varphi(\boldsymbol{x}, \cdot) \in \mathbb{R}$ *ako výslednú distribučnú funkciu straty* $z = f(\boldsymbol{x},\boldsymbol{y})$*, pre ktorú platí*

$$\varphi(\boldsymbol{x},\zeta) = P[\boldsymbol{y}|f(\boldsymbol{x},\boldsymbol{y}) \leq \zeta].$$

1.2 Zavedenie mier rizika VaR a CVaR

Strata v našom prípade predstavuje určitú formu rizika. V praxi sa preto často stretávame s jeho klasifikáciou pomocou hodnôt VaR (Value at Risk) a $CVaR$ (Conditional Value at Risk), prekladané ako hodnota v riziku a podmienená hodnota v riziku. Jednotlivé hodnoty udávajú, aká je možná výška straty na určitej hladine spoľahlivosti α (väčšinou volenej 0.99, 0.95) za dané časové obdobie (napr. 1 deň, 1 rok). Tieto hodnoty si následne zadefinujeme na základe definícií uvedených v časti 1.1.

Definícia 4 (VaR). *(Rockafellar a Uryasev (2002), str. 1447). Nech* $\alpha - VaR$ *straty vzhľadom k rozhodovaciemu vektoru* $\boldsymbol{x}$ *je hodnota*

$$\zeta_\alpha(\boldsymbol{x}) = \min\{\zeta \in \mathbb{R}|\varphi(\boldsymbol{x},\zeta) \geq \alpha\}, \tag{1.1}$$

kde v 1.1 je dosiahnuté minimum, keďže funkcia $\varphi(\boldsymbol{x},\zeta)$ *je nerastúca a zprava spojitá v* ζ*. Ak* $\varphi(\boldsymbol{x},\cdot)$ *je spojitá a striktne rastúca,* $\zeta_\alpha(\boldsymbol{x})$ *je jednoznačne určené* ζ *spĺňajúce* $\varphi(\boldsymbol{x},\zeta) = \alpha$*. Inak riešenie rovnice* $\varphi(\boldsymbol{x},\zeta) = \alpha$ *neexistuje, alebo ich existuje nespočetné mnoho.*

Definícia 5 ($CVaR$). *(Rockafellar a Uryasev (2002),str.1448). Predpokladajme, že* $E[|f(\boldsymbol{x},\boldsymbol{y})|] < \infty$ *pre* $\forall \boldsymbol{x} \in X$*, potom* $\alpha - CVaR$ *straty vzhľadom k rozhodovaciemu vektoru* $\boldsymbol{x}$ *je hodnota*

$$\phi_\alpha(\boldsymbol{x}) = \textit{strednej hodnote } \alpha - \textit{chvosta rozdelenia } f(\boldsymbol{x},\boldsymbol{y}),$$

kde ide o rozdelenie dané distribučnou funkciou $\varphi_\alpha(\boldsymbol{x},\cdot)$ *definovanou nasledovne*

$$\varphi_\alpha(\boldsymbol{x},\zeta) = \begin{cases} 0 & \textit{pre } \zeta < \zeta_\alpha(\boldsymbol{x}), \\ [\varphi(\boldsymbol{x},\zeta) - \alpha]/[1-\alpha] & \textit{pre } \zeta \geq \zeta_\alpha(\boldsymbol{x}). \end{cases}$$

Vzhľadom k hodnote $\alpha \in (0,1)$ *predstavujúcej danú hladinu spoľahlivosti.*

Je dobré poznamenať, že distribučná funkcia $\varphi_\alpha(\boldsymbol{x},\cdot)$, je iná distribučná funkcia než zmienená funkcia v definícií 3, avšak podobne ako $\varphi(\boldsymbol{x},\cdot)$ je neklesajúca, zprava spojitá, kde $\varphi_\alpha(\boldsymbol{x},\zeta) \to 1$ pre $\zeta \to \infty$. Pomocou tejto distribučnej funkcie vieme však dobre definovať rozdelenie uvažovaného α-chvostu.

Vzhľadom na definície 4 a 5 môžme medzi jednotlivými hodnotami pozorovať rôzne rozdiely. Jedným z hlavných rozdielov je, že hodnota VaR nám dáva číslo, ktoré zodpovedá najmenšej hodnote straty, ktorú s uvedenou hladinou spoľahlivosti α dané portfólio neprekročí.

Na druhej strane, hodnota $CVaR$ hovorí o tom, aká veľká bude očakávaná strata, ak danú hodnotu VaR strata prekročí, čo je pre nás omnoho smerodajnejšia informácia. V následujúcej podkapitole si uvedieme niektoré z ďalších charakteristík týchto hodnôt.

1.3 Základné charakteristiky pre VaR a CVaR

V tejto časti sa budeme bližšie zaoberať mierami rizika VaR a $CVaR$. Zameriame sa na ich vlastnosti a taktiež budeme skúmať ich vzájomné súvislosti, s ktorými budeme ďalej pracovať.

Veta 1 (Rockafellar a Uryasev (2002), str.1452)**.** *Nech $\lambda_\alpha(\boldsymbol{x})$ je pravdepodobnosť toho, že strata $z = \zeta_\alpha(\boldsymbol{x})$, ktorá je daná rozdelením α-chvostu uvedeného v definícií 5, konkrétne*

$$\lambda_\alpha(\boldsymbol{x}) = [\varphi(\boldsymbol{x},\zeta_\alpha(\boldsymbol{x})) - \alpha]/[1 - \alpha] \in [0,1].$$

Potom môže nastať jedna z nasledujúcich možností:

- $\varphi(\boldsymbol{x},\zeta_\alpha(\boldsymbol{x})) < 1$*, teda existuje šanca, že strata bude vyššia ako* $\zeta_\alpha(\boldsymbol{x})$ *a*

$$\phi_\alpha(\boldsymbol{x}) = \lambda_\alpha(\boldsymbol{x})\zeta_\alpha(\boldsymbol{x}) + [1 - \lambda_\alpha(\boldsymbol{x})]\phi_\alpha^+(\boldsymbol{x}),$$

  kde $\phi_\alpha^+(\boldsymbol{x}) = \mathsf{E}[f(\boldsymbol{x},\boldsymbol{y})|f(\boldsymbol{x},\boldsymbol{y}) > \zeta_\alpha(\boldsymbol{x})]$.

- $\varphi(\boldsymbol{x},\zeta_\alpha(\boldsymbol{x})) = 1$*, tým pádom* $\zeta_\alpha(\boldsymbol{x})$ *je najvyššia možná strata, ktorá môže nastať, z kadial vyplýva rovnosť*

$$\phi_\alpha(\boldsymbol{x}) = \zeta_\alpha(\boldsymbol{x}).$$

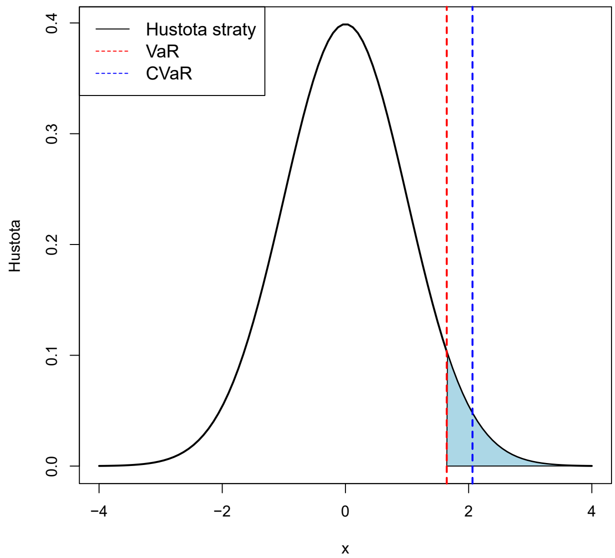
Dôkaz. Tieto vzťahy jednoznačne vyplývajú z definície 5 a z platnosti nerovnosti $\alpha \leq \varphi(\boldsymbol{x},\zeta_\alpha(\boldsymbol{x}))$, ktorá vďaka definícií 4 platí vždy.

❏

Dôsledok (Rockafellar a Uryasev (2002), str. 1452). Na základe definícií 4 ,5 a vety 1 platí

$$\phi_\alpha(\boldsymbol{x}) \geq \zeta_\alpha(\boldsymbol{x}).$$

Dokonca, v prípade kedy neexistuje šanca, že strata bude väčšia ako $\zeta_\alpha(\boldsymbol{x})$ platí $\phi_\alpha(\boldsymbol{x}) = \zeta_\alpha(\boldsymbol{x})$.



Obr. 1.1: Hodnoty VaR a $CVaR$ v prípade predpokladu normovaného normálneho rozdelenia straty na hladine $\alpha = 0.95$.

Veta 2 (Rockafellar a Uryasev (2002), str. 1453)**.** *Predpokladajme, že pravdepodobnostná miera P je koncentrovaná v konečne mnoho bodoch $y_k \in Y$ tak, že pre každé $\boldsymbol{x} \in X$ distribúcia straty $z = f(\boldsymbol{x},\boldsymbol{y})$ je podobne koncentrovaná v konečne mnoho bodoch a $\varphi(\boldsymbol{x},\cdot)$ je po častiach lineárna funkcia so skokmi v jednotlivých bodoch. Pre pevné $\boldsymbol{x}$, nech sú zodpovedajúce realizácie straty usporiadané nasledovne $z_1 < z_2 < \cdots < z_N$, kde $P(z = z_k) = p_k$ a $p_k > 0$ pre $k = 1, \ldots, N$. Nech k_α je jednoznačne určený index spĺňajúci*

$$\sum_{k=1}^{k_\alpha} p_k \geq \alpha > \sum_{k=1}^{k_\alpha - 1} p_k.$$

Potom pre $\alpha - VaR$ straty platí

$$\zeta_\alpha(\boldsymbol{x}) = z_{k_\alpha},$$

a $\alpha - CVaR$ dostávame z rovnosti

$$\phi_\alpha(\boldsymbol{x}) = \frac{1}{1-\alpha}\Big[\Big(\sum_{k=1}^{k_\alpha} p_k - \alpha\Big) z_{k_\alpha} + \sum_{k=k_\alpha+1}^{N} p_k z_k\Big].$$

Dôkaz. Jednotlivé kroky dôkazu nájdeme v článku (Rockafellar a Uryasev (2002), str. 1453).

❑

Kapitola 2

Reformulácia mier v riziku

V tejto časti bude našim cieľom vyjadriť hodnoty VaR a $CVaR$ pomocou jednej funkcie, ktorej konštrukciu postavíme na základe definícií a viet, ktoré sme si uviedli v prvej kapitole. Zadefinovanie takejto funkcie nám neskôr prinesie možnosti na jednoduchší výpočet hodnôt v riziku.

2.1 Konštrukcia funkcie pre VaR a CVaR

Z definície 4 a za predpokladu existencie riešenia rovnosti $\varphi(\boldsymbol{x},\zeta) = \alpha$ dostávame, že $\zeta_\alpha(\boldsymbol{x})$ predstavuje ľavý krajný bod neprázdneho intervalu, ktorý obsahuje hodnoty ζ také, že $\varphi(\boldsymbol{x},\zeta) = \alpha$. Rovnosť plynie z vlastností distribučnej funkcie $\varphi(\boldsymbol{x},\zeta)$, ktorá je spojitá a neklesajúca vzhľadom k premennej ζ. Na druhej strane z definície 5 dostávame, že $P[f(\boldsymbol{x},\boldsymbol{y}) \geq \zeta_\alpha(\boldsymbol{x})] = 1 - \alpha$. Teda hodnota $\phi_\alpha(\boldsymbol{x})$ zodpovedá podmienenej strednej hodnote straty vzhľadom k $\boldsymbol{x}$ ku ktorému bude strata rovná hodnote $\zeta_\alpha(\boldsymbol{x})$ alebo väčšia.

Na základe týchto poznatkov a predpokladu, že náhodný vektor $\boldsymbol{y}$ definovaný na pravdepodobnostnom priestore $(Y, \mathcal{B}, P)$ pochádza zo spojitého rozdelenia definovaného hustotou $p(\boldsymbol{y})$, môžme charakterizovať hodnoty $\zeta_\alpha(\boldsymbol{x})$ a $\phi_\alpha(\boldsymbol{x})$ pomocou funkcie F_α definovanej na množine $X \times \mathbb{R}$ následovne

$$F_\alpha(\boldsymbol{x},\zeta) = \zeta + \frac{1}{1-\alpha} \int_{\boldsymbol{y}\in\mathbb{R}^n} [f(\boldsymbol{x},\boldsymbol{y}) - \zeta]^+ p(\boldsymbol{y}) \mathrm{d}y, \tag{2.1}$$

kde $[t]^+ = \max\{0,t\}$ pre $t = [f(\boldsymbol{x},\boldsymbol{y}) - \zeta]$.

Jednou z najdôležitejších vlastností funkcie F_α definovanou pomocou 2.1 je jej konvexita, ktorá je kľúčovou z pohľadu optimalizácie. Keďže našim cieľom bude funkciu neskôr minimalizovať vzhľadom k premennej ζ. Na túto vlastnosť a jej dôsledky sa bližšie pozrieme v následujúcej vete 3.

Veta 3. *(Rockafellar a Uryasev (2000), str.5) Nech náhodný vektor $\boldsymbol{y}$ definovaný na pravdepodobnostnom priestore $(Y, \mathcal{B}, P)$ pochádza zo spojitého rozdelenia definovaného hustotou $p(\boldsymbol{y})$. Potom odvodená funkcia $F_\alpha(\boldsymbol{x},\zeta)$ je ako funkcia premennej $\zeta \in \mathbb{R}$ konvexná a spojite diferencovateľná. Hodnotu $\alpha - CVaR$ straty vzhľadom k ľubovolnému vektoru $\boldsymbol{x}$ vieme určiť z rovnosti*

$$\phi_\alpha(\boldsymbol{x}) = \min_\zeta F_\alpha(\boldsymbol{x},\zeta). \tag{2.2}$$

Množinu hodnôt ζ, pre ktoré funkcia nadobúda minimum označíme

$$A_\alpha(\boldsymbol{x}) = \arg\min_{\zeta} \; F_\alpha(\boldsymbol{x},\zeta), \tag{2.3}$$

ktorá je neprázdnym, uzavretým a obmedzeným intervalom (väčšinou zredukovaným na jeden bod) a hodnota $\alpha - VaR$ straty je daná ako

$$\zeta_\alpha(\boldsymbol{x}) = \textit{ľavý krajný bod intervalu} \;\; A_\alpha(\boldsymbol{x}).$$

Konkrétne vždy platí

$$\zeta_\alpha(\boldsymbol{x}) \in \arg\min_{\zeta} \; F_\alpha(\boldsymbol{x},\zeta) \quad a \quad \phi_\alpha(\boldsymbol{x}) = F_\alpha(\boldsymbol{x},\zeta_\alpha(\boldsymbol{x})).$$

Dôkaz. Jednotlivé kroky dôkazu sú podrobne popísané v článku Rockafellar a Uryasev (2000) str. 25.

❑

Vetou 3 sa autori zaberali aj neskôr v článku Rockafellar a Uryasev (2002), kde môžme nájsť jej formuláciu pre ľubovolné rozdelenie.

Ďalej poznamenajme, že z hľadiska výpočtu vieme rovnako dobre minimalizovať funkciu $(1-\alpha)F_\alpha(\boldsymbol{x},\zeta)$ ako aj $F_\alpha(\boldsymbol{x},\zeta)$. Týmto spôsobom sa vieme vyhnúť deleniu integrálu výrazom $(1-\alpha)$, čo môže byť z numerického hľadiska výhodnejšie najmä v prípadoch, kedy je hodnota $(1-\alpha)$ príliš malá.

Na základe vety 3, ktorá nám poskytuje informáciu o vlastnostiach funkcie $F_\alpha(\boldsymbol{x},\zeta)$ a vedomostiam z teórie optimalizácie vieme, že funkcie ktoré sú spojito diferencovateľné a konvexné sa dajú jednoducho numericky minimalizovať. Práve konvexita, ktorá je kľúčovou vlastnosťou z hľadiska optimalizácie nám zaručí, že pri postupnom prechádzaní funkcie a hľadaním minima nenatrafíme na lokálne minimum, ktoré by bolo rozdielne od globálneho. To znamená, že optimálna hodnota je určená jednoznačne.

Rovnosť 2.2 nám poskytuje prístup k výpočtu hodnoty $\alpha - CVaR$ bez nutnosti prvotného výpočtu hodnoty $\alpha - VaR$ na základe ktorej je podmienená hodnota v riziku na danej úrovni spoľahlivosti α definovaná (viz. Definícia 5). Avšak v prípade potreby vieme hodnotu $\alpha - VaR$ získať ako vedľajší produkt, ktorý získame určením intervalu $A_\alpha(\boldsymbol{x})$ definovaného pomocou 2.3 a extrahovaním jeho ľavého krajného bodu, pokiaľ daný interval obsahuje viac ako jeden bod.

Pre jednoduchšie narábanie s funkciou $F_\alpha(\boldsymbol{x},\zeta)$, vieme jej časť, ktorá je definovaná pomocou integrálu (2.1) nahradiť jeho príslušnou aproximáciou.

Jednou z možností je, že si vygenerujeme náhodný výber vektorov $\boldsymbol{y_1},\boldsymbol{y_2},\ldots,\boldsymbol{y_T}$ z rozdelenia daného hustotou $p(\boldsymbol{y})$ vzhľadom ku ktorej bude aproximácia funkcie $F_\alpha(\boldsymbol{x},\zeta)$ vyzerať nasledovne

$$\tilde{F}_\alpha(\boldsymbol{x},\zeta) = \zeta + \frac{1}{(1-\alpha)T} \sum_{t=1}^{T} \left[f(\boldsymbol{x},\boldsymbol{y}_t) - \zeta \right]^+. \tag{2.4}$$

Funkcia $\tilde{F}_\alpha(\boldsymbol{x},\zeta)$ je konvexná a po častiach lineárna vzhľadom k ζ, avšak nieje diferencovateľná vzhľadom k ζ. Napriek tomu ju však môžme minimalizovať pomocou riešenia problému elementárneho lineárneho programovania.

Nasledujúca veta, ktorú si zavedieme opäť poukazuje na dôležitosť vyjadrenia hodnôt $\alpha - VaR$ a $\alpha - CVaR$ pomocou funkcie F_α. Na základe jej vlastností uvedených vo vete 3 sa dostávame k formálnemu vyjadreniu riešenia minimalizačného problému.

Veta 4. *(Rockafellar a Uryasev (2000), str. 6) Predpokladajme, že náhodný vektor* $\boldsymbol{y}$ *definovaný na pravdepodobnostnom priestore* $(Y, \mathcal{B}, P)$ *pochádza zo spojitého rozdelenia definovaného hustotou* $p(\boldsymbol{y})$. *Minimalizácia hodnoty* $\alpha - CVaR$ *straty vzhľadom k rozhodovaciemu vektoru* $\boldsymbol{x} \in X$ *je ekvivalentná minimalizácií funkcie* $F_\alpha(\boldsymbol{x},\zeta)$ *cez všetky* $(\boldsymbol{x},\zeta) \in X \times \mathbb{R}$, *v zmysle*

$$\min_{\boldsymbol{x}\in X} \phi_\alpha(\boldsymbol{x}) = \min_{(\boldsymbol{x},\zeta)\in X\times\mathbb{R}} F_\alpha(\boldsymbol{x},\zeta),$$

kde naviac pre dvojicu $(\boldsymbol{x}^*,\zeta^*)$ *dosahuje funkcia na pravej strane rovnosti hodnotu minima, práve vtedy a len vtedy ak aj funkcia na ľavej strane rovnosti dosahuje minimum pre hodnotu* $\boldsymbol{x}^*$ *a* $\zeta^* \in A_\alpha(\boldsymbol{x}^*)$. *Najmä preto, za okolností, keď je interval* $A_\alpha(\boldsymbol{x}^*)$ *zredukovaný na jeden bod (ako bežne býva), minimalizáciou funkcie* $F_\alpha(\boldsymbol{x},\zeta)$ *cez všetky* $(\boldsymbol{x},\zeta) \in X \times \mathbb{R}$ *dostávame ako riešenie dvojicu* $(\boldsymbol{x}^*,\zeta^*)$, *nie nutne jednoznačnú, takú že* $\boldsymbol{x}^*$ *minimalizuje hodnotu* $\alpha - CVaR$ *a zároveň* ζ^* *zodpovedá korešpondujúcej hodnote* $\alpha - VaR$.

Dôkaz. Jednotlivé kroky dôkazu nájdeme v článku Rockafellar a Uryasev (2000), uvedené na strane 26.

❑

Veta 5. *(Rockafellar a Uryasev (2002), str. 1457) Ak je funkcia* $f(\boldsymbol{x},\boldsymbol{y})$ *konvexná vzhľadom k premennej* $\boldsymbol{x}$ *a pre každé* $\boldsymbol{y}$ *potom,* $\phi_\alpha(\boldsymbol{x})$ *je taktiež konvexná vzhľadom k premennej* $\boldsymbol{x}$ *a podobne aj funkcia* $F_\alpha(\boldsymbol{x},\zeta)$ *je konvexná vzhľadom k dvojici* $(\boldsymbol{x},\zeta)$.

Dôkaz. Podrobné prevedenie dôkazu nájdeme v článku Rockafellar a Uryasev (2002) na strane 1457.

❑

Dôsledok. Ak sú podmienky minimalizácie funkcie v tvare, ktorý spĺňa podmienky vety 5, smerujeme k riešeniu minimalizačnej konvexnej úlohy, keďže aj množina X je konvexnou množinou.

2.2 Zhrnutie vlastností funkcie

Veta 4 nám poskytuje možnosť, ako jednoduchým spôsobom vyčíslime hodnotu zodpovedajúcu minimu $\alpha - CVaR$ a hodnotu $\alpha - VaR$ zároveň. Pri výpočte $\boldsymbol{x}$, ktoré zodpovedá hodnote minima $\alpha - CVaR$, nieje teda nutné pracovať priamo s funkciou $\phi_\alpha(\boldsymbol{x})$, ale môžme použiť odvodenú funkciu $F_\alpha(\boldsymbol{x},\zeta)$.

Naďalej budeme pre jednoduchosť pracovať s funkciou $F_\alpha(\boldsymbol{x},\zeta)$, ktorá je podľa vety 5 konvexná vzhľadom k premennej ζ a taktiež k dvojici $(\boldsymbol{x},\zeta)$. Zároveň si môžme povšimnúť, že z konvexity funkcie $f(\boldsymbol{x},\boldsymbol{y})$ vzhľadom k $\boldsymbol{x}$ vyplývala aj konvexita aproximačnej funkcie $\tilde{F}_\alpha(\boldsymbol{x},\zeta)$ taktiež vzhľadom k dvojici $(\boldsymbol{x},\zeta)$, kde funkcia $\tilde{F}_\alpha(\boldsymbol{x},\zeta)$ je zadefinovaná pomocou rovnosti 2.4.

Minimalizačný prístup pre hodnotu $\phi_\alpha(\boldsymbol{x})$ uvedený vo vete 4 môžme využiť v kombinácií s aproximačnou metódou integrálu v definícií funkcie $F_\alpha(\boldsymbol{x},\zeta)$.

Dostávame sa teda k tomu, že problém minimalizácie funkcie F_α na množine $X \times \mathbb{R}$ spadá do kategórie konvexného programovania. Z toho dôvodu existuje široká škála spôsobov ako danú problematiku hľadania minima riešiť, kde práve veta 4 nám dovoľuje použiť tieto techniky pri minimalizácií hodnoty $\alpha - CVaR$.

Kapitola 3

Konštrukcia lineárneho programu pre portfólio

V tejto časti sa budeme zaoberať odvodením lineárneho programu pre optimalizáciu portfólia. Našim cieľom bude, aby naše portfólio čo najlepšie replikovalo zvolený index, teda aby odchýlka medzi cenami portfólia a uvažovaného indexu bola čo najmenšia. Riešenie tohto problému vedie na úlohu konvexného programovania, ktorú prevedieme na lineárnu z dôvodu jednoduchosti jej výpočtu. Ďalej uvedieme využitie reformulácie pre hodnoty $CVaR$ a VaR pomocou funkcie F_α, ktorú použijeme ako jedno z obmedzení pri riešení minimalizačnej úlohy.

3.1 Značenie

Uvažujeme index I, pozostávajúci z viacerých akcií. Na základe tohto indexu si vyberieme konkrétne akcie S_j, $j = 1,\dots,n$, ktoré sú súčasťou indexu I. Označíme I_t cenu indexu I v čase t pre $t = 1,\dots,T$ a $y_{t,j}$ cenu akcie S_j v čase t. Nech ν je hodnota investície do portfólia v čase 1. Potom označíme $\theta = \nu/I_1$ počet jednotiek indexu I v čase 1. Nech x_j, pre $j = 1,\dots,n$ je počet jednotiek akcie S_j v navrhovanom portfóliu, pomocou ktorého budeme index replikovať. Na základe zavedeného označenia dostávame, že hodnota portfólia v čase t je rovná $\sum_{j=1}^{n} y_{t,j}x_j$.

Následne vieme povedať, že absolútna hodnota relatívnej odchýlky hodnoty replikovaného portfólia od cieľovej hodnoty $\nu = \theta I_t$ je

$$\big|(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j)/\theta I_t\big|.$$

Je dobré si uvedomiť, že absolútna hodnota relatívnej odchýlky nám udáva, ako veľmi je vzdialené portfólio od daného indexu, či už smerom nahor alebo nadol.

Uvažovaný vektor cien $\boldsymbol{y}_t = (y_{t,1},\dots,y_{t,n})$ pre $t = 1,\dots,T$, získame z pozorovaní náhodného vektoru $\boldsymbol{y} \in \mathbb{R}^n$, zavedeného pri definícií straty 2. Následne predpokladáme, že náhodný vektor $\boldsymbol{y}$ má diskrétne rovnomerné rozdelenie na množine pozorovaní $\{\boldsymbol{y}_1,\dots,\boldsymbol{y}_T\}$ a teda platí $P[\boldsymbol{y} = \boldsymbol{y}_t] = 1/T$ pre všetky t.

Vezmeme funkciu straty $f(\boldsymbol{x},\boldsymbol{y})$, pomocou ktorej zavedieme relatívnu hodnotu straty. V našom prípade relatívna strata predstavuje spomínanú relatívnu odchýlku nášho portfólia od indexu, pomocou ktorého chceme uvažovaný index replikovať, teda

$$f(\boldsymbol{x},\boldsymbol{y_t}) = \big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\big)/\theta I_t. \tag{3.1}$$

V tomto zápise funkciu následne využijeme ako funkciu pomocou ktorej budeme minimalizovať strednú hodnotu absolútnej odchýlky $|f(\boldsymbol{x},\boldsymbol{y_t})|$.

3.2 Formulácia optimalizačnej úlohy

K nájdeniu optimálneho riešenia sa formálne dostávame k minimalizačnej úlohe, pomocou ktorej budeme naše portfólio replikovať. Ako účelovú funkciu zvolíme absolútnu hodnotu funkcie $f(\boldsymbol{x},\boldsymbol{y_t})$ zadefinovanú v 3.1, ktorú budeme minimalizovať v jednotlivých časových intervaloch $1,\ldots,T$ tak, aby súčet výsledných absolútnych odchýliek bol čo najmenší. Optimalizačnú úlohu formulujeme v tvare

$$min\ g(\boldsymbol{x}) = \frac{1}{T}\sum_{t=1}^{T}\Big|\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\Big|, \tag{3.2}$$

za podmienok :

$$\sum_{j=1}^{n} y_{1,j}x_j = \nu \tag{3.3}$$

$$\zeta + \frac{1}{(1-\alpha)T}\sum_{t=1}^{T}\big[f(\boldsymbol{x},\boldsymbol{y_t}) - \zeta\big]^+ \leq \omega \tag{3.4}$$

$$x_j \geq 0,\ j = 1,\ldots,n \tag{3.5}$$

kde minimalizujeme funkciu vzhľadom k vektoru $\boldsymbol{x} = (x_1,\ldots,x_n)$ a premennej ζ. Podmienka 3.3 nám zabezpečí, že portfólio bude zostavené, tak aby sme investovali celú počiatočnú investíciu ν. Teda, na základe cien jednotlivých akcií v čase 1 sa rozhodneme koľko jednotiek danej akcie kúpime, aby celkový súčet investícií bol v hodnote ν. Číslo ω v podmienke 3.4 je maximálna tolerovaná hodnota pre $CVaR$, ktorú volíme tak, aby minimalizačná úloha bola prípustná. Funkcia straty $f(\boldsymbol{x},\boldsymbol{y_t}) = (\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j)/\theta I_t$, tým pádom vieme podmienku 3.4 ekvivalentne zapísať do tvaru

$$\zeta + \frac{1}{(1-\alpha)T}\sum_{t=1}^{T}\Big[\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t - \zeta\Big]^+ \leq \omega. \tag{3.6}$$

Výraz na ľavej strane nerovnosti (3.6) je rovný funkcií $F_\alpha(\boldsymbol{x},\zeta)$ z definície 2.1.V diskrétnom prípade, ktorý uvažujeme, aplikáciou operátora strednej hodnoty E je

$$E[f(\boldsymbol{x},\boldsymbol{y}) - \zeta]^+ = \frac{1}{T}\sum_{t=1}^{T}\big[[\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j)/\theta I_t] - \zeta\big]^+,$$

z kadiaľ sa odvíja rovnosť pravej strany s funkciou.

Problémom funkcie $g(\boldsymbol{x})$ vzhľadom k výpočtu je, že nie je lineárna kvôli absolútnej hodnote v jej predpise. Avšak pre ľubovolné α a ω, vieme tento problém previezť na úlohu lineárneho programovania.

3.3 Prevedenie nelineárneho programu na lineárny

Ako prvé si pri linearizácií daného programu označíme problémovú zložku funkcie $g(\boldsymbol{x})$. V našom prípade ide o absolútnu hodnotu, ktorá porušuje uvažovanú funkciu. Nech $\eta_{t,0} = |f(\boldsymbol{x},\boldsymbol{y}_t)| = |(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j)/\theta I_t|$ pre všetky $t = 1,\ldots,T$. So zavedením ďalších podmienok môžeme úlohu formulovať v tvare

$$min\ g(\boldsymbol{x}) = \frac{1}{T}\sum_{t=1}^{T}\eta_{t,0} \tag{3.7}$$

za podmienok :

$$\eta_{t,0} = \left|\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\right|,\ t = 1,\ldots,T \tag{3.8}$$

a podmienok 3.3, 3.5 a 3.6 uvedených vyššie.Vzhľadom na to, že minimalizujeme účelovú funkciu, môžeme v podmienke 3.8 namiesto rovnosti použiť nerovnosť $\geq$, čím sa nám program nezmení keďže pre

$$\eta_{t,0} \geq \left|\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\right|,$$

platí

$$\left|\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\right| = \max\left[\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\ ,-\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\right] \tag{3.9}$$

$$\eta_{t,0} \geq \max\left[\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\ ,-\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\right], \tag{3.10}$$

$$\Leftrightarrow\ \eta_{t,0} \geq \Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t\ ,\ \eta_{t,0} \geq -\Big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\Big)/\theta I_t. \tag{3.11}$$

Do rozporu s linearitou sa však dostáva ešte podmienka 3.6, kde uvažujeme iba kladnú časť. Na vyriešenie tohto problému si zavedieme pomocnú premennú u_t pomocou ktorej podmienku linearizujeme.

Náš finálny lineárny program, ktorý bude riešiť problematiku replikácie indexu pomocou nášho portfólia pre ľubovolné hodnoty α a ω na základe odvodení vyššie, bude potom v tvare

$$min\ g(\boldsymbol{x}) = \frac{1}{T}\sum_{t=1}^{T}\eta_{t,0}$$

za podmienok :

$$\sum_{j=1}^{n} y_{1,j}x_j = \nu$$

$$\big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\big)/\theta I_t \leq \eta_{t,0},\ t = 1,\ldots,T$$

$$-\ \big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\big)/\theta I_t \leq \eta_{t,0},\ \ t = 1,\ldots,T$$

$$\zeta + \frac{1}{(1-\alpha)T}\sum_{t=1}^{T} u_t \leq \omega$$

$$\big[\big(\theta I_t - \sum_{j=1}^{n} y_{t,j}x_j\big)/\theta I_t - \zeta\big] \leq u_t,\ t = 1,\ldots T$$

$$u_t \geq 0,\ \ t = 1,\ldots,T$$

$$\eta_{t,0} \geq 0,\ \ t = 1,\ldots,T$$

$$x_j \geq 0,\ j = 1,\ldots,n$$

ktorý budeme minimalizovať vzhľadom k $\boldsymbol{x} = (x_1,\ldots,x_n)$, premennej ζ a k novozavedeným premenným η a u.

Niektorými časťami minimalizačnej úlohy sa zaoberali aj autori v článku Rockafellar a Uryasev (2002), kde linearizovali podmienku 3.8 pomocou zavedenia pomocnej premennej η a jej rozkladu na kladnú a zápornú časť. My sme sa týmto rozkladom podrobnejšie zaoberali v rovniciach 3.9, 3.10 a 3.11. Následne sme navyše zaviedli pomocnú premennú u, pomocou ktorej sme previedli podmienku 3.6 do lineárneho tvaru. Tieto kroky nám zaručili, že účelová funkcia a aj všetky jej podmienky minimalizácie budú v lineárnom tvare, čo znamená, že môžme úlohu riešiť pomocou nástrojov lineárneho programovania.

Kapitola 4

Praktická aplikácia

Finálny tvar optimalizačnej úlohy, ktorú sme odvodili na konci kapitoly 3.3 využijeme pri replikácií indexu S&P 500 pomocou nami zvolených akcií v portfóliu. Na riešenie minimalizačnej úlohy využijeme programovací jazyk Python a balíček PuLP, kde následne prevedieme výpočet v dvoch fázach. Na konkrétnu konštrukciu tohto kódu je možné nahliadnuť v prílohe tejto práce.

V prvej fáze, takzvanej $in-sample$, prevedieme výpočet optimalizačnej úlohy na základe dostupných historických dát. V druhej fáze $out-of-sample$ aplikujeme výsledky z prvej fázy na dáta, ktoré neboli v kategórií $in-sample$ a budeme pozorovať nakoľko presnú replikáciu sme previedli.

4.1 Predstavenie problému

Portfólio, pomocou ktorého budeme replikovať spomínaný index $I =$ S&P 500, bude pozostávať z nasledujúcich dvadsiatich akcii, ktorých názov je uvedený pod takzvanými tickermi: AAPL, WMT, IBM, MU, BA, AXP, META, PEP, MCD, NFLX, ADI, BAC, PFE, KO, MA, TSLA, HPQ, STT, CMG, PSX. V zmysle zavedeného značenia v podkapitole 3.1 akcie predstavujú jednotlivé zložky S_j, kde $j = 1, \ldots ,20$.

Ďalej budeme pri výpočte uvažovať

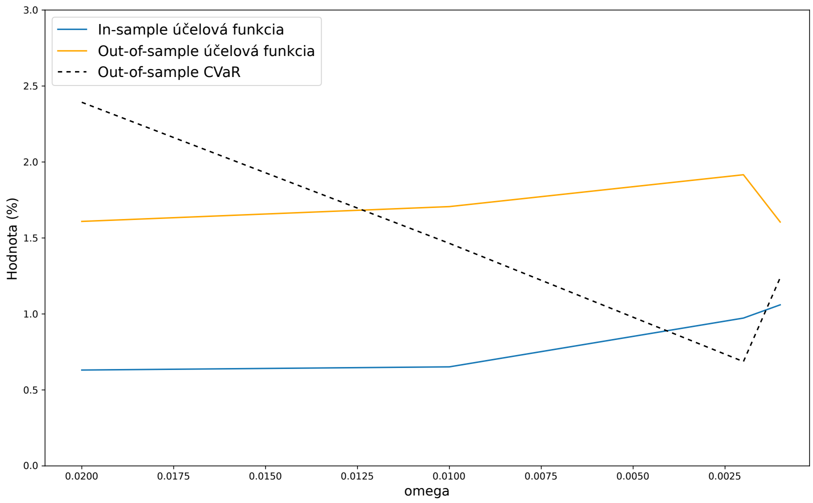
- Časové obdobie od 08.06.2020 do 12.12.2022 predstavujúce 131 týždňov, kde do fázy $in-sample$ zaradíme týždne v časovom období od 08.06.2020 do 06.06.2022 a do fázy $out-of-sample$ týždne od 06.06.2022 do 12.12.2022
- Pondelkové uzatváracie ceny jednotlivých aktív, upravené o trhové parametre v týždňoch $t = 1, \ldots ,131$
- Cenu indexu I_t v týždni $t = 1, \ldots ,131$
- Náhodný vektor cien $y_{t,j}$, ktorý predstavuje cenu akcie S_j pre $j = 1, \ldots ,20$ v danom týždni $t = 1, \ldots ,131$
- Počiatočnú investíciu do portfólia v hodnote $\nu = 1000$ USD
- Počet jednotiek $\theta = \nu/I_1$ indexu S&P 500 vzhľadom k počiatočnej investícii
- Hodnotu x_j predstavujúcu počet jednotiek akcie S_j pre $j = 1, \ldots ,20$ v portfóliu

- Úroveň spoľahlivosti $\alpha = 0.95$ pre mieru rizika $CVaR$
- Toleranciu podmienenej hodnoty v riziku ω

Ako prvý prevedieme výpočet $in - sample$, ktorý prebieha v týždňoch $t = 1, \ldots ,104$ pre rôzne hodnoty ω, na úrovni spoľahlivosti α. Následne použijeme riešenia k výpočtu hodnoty účelovej funkcie a miery rizika $CVaR$ vo fáze $out - of - sample$ v týždňoch $t = 105, \ldots , 131$. Na základe výsledných hodnôt, ktoré sú uvedené tabuľke 4.1, môžme pozorovať, že pri hodnote $\omega = 0.02$ dosahuje účelová funkcia v období $in - sample$ najnižšiu hodnotu. Znamená to teda, že odchýlky medzi replikačným portfóliom a indexom sú najmenšie. Z toho dôvodu si na prevedenie výpočtu optimalizačnej úlohy zvolíme pevnú hodnotu $\omega = 0.02$.

Hodnota ω	Hodnota účelovej funkcie vo fáze $in - sample$ (%)	Hodnota účelovej funkcie vo fáze $out - of - sample$ (%)	Hodnota $CVaR$ vo fáze $out - of - sample$ (%)
0.02	0.63083	1.60902	2.3930
0.01	0.65183	1.70646	1.4646
0.002	0.97288	1.91618	0.6865
0.001	1.05997	1.60507	1.2386

Tabuľka 4.1: Hodnoty funkcií v závislosti na premennej ω.



Obr. 4.1: Priebeh jednotlivých funkcií v závislosti na hodnotách ω uvedených v tabuľke 4.1.

Z obrázku 4.1 môžme vidieť, že zvolená hodnota ω má najväčší vplyv na mieru rizika $CVaR$ v období $out - of - sample$, kedy do určitého bodu jej miera spoločne s klesajúcou hodnotou ω taktiež klesá. Čo sa týka hodnôt účelovej funkcie, obe so zmenšujúcou sa hodnotou ω mierne stúpajú, čo sme aj očakávali, keďže zmenšovaním hodnoty ω kladieme vyššie nároky na podmienku, čím sa stáva prísnejšou, čoho následkom je nárast veľkosti odchýliek. Avšak na konci môžeme pozorovať výraznejší pokles účelovej funkcie v období $out - of - sample$, ale naopak nárast miery rizaka $CVaR$ pre dané obdobie.

4.2 Prevedenie optimalizačnej úlohy

Začíname opäť výpočtom vo fáze $in - sample$ pre hodnoty uvedené v podkapitole 4.1 a pre pevnú hodnotu $\omega = 0.02$. Účelová funkcia v tejto fáze pre dané hodnoty spĺňa nasledujúcu rovnosť

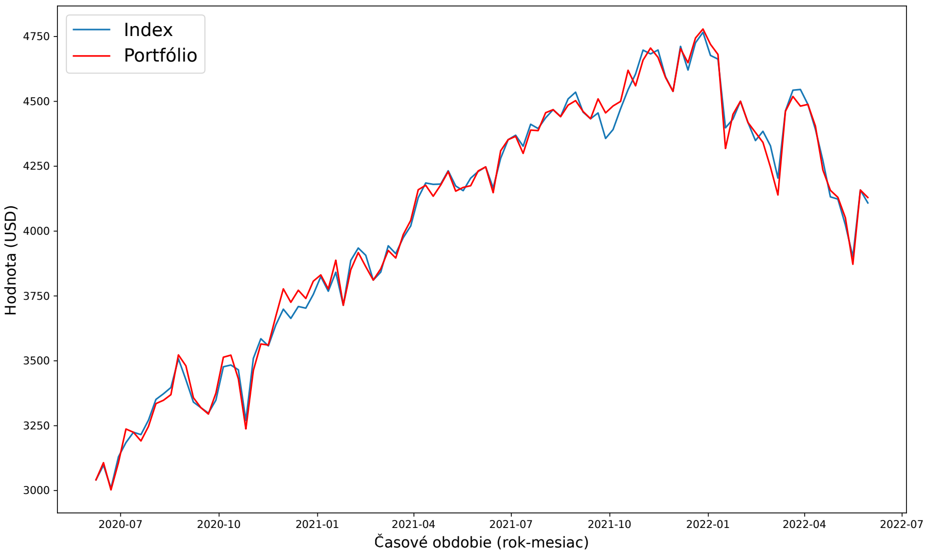
$$min\ g(\boldsymbol{x}) = \frac{1}{104}\sum_{t=1}^{104}\eta_{t,0} = 0.0063083.$$

Zároveň pre optimálne hodnoty $\boldsymbol{x}^* = (x_1^*, \ldots, x_{20}^*)$ predstavujúce počet jednotiek jednotlivých akcií v našom portfóliu platí následujúca tabuľka 4.2.

Ticker	Premenná	Počet jednotiek
AAPL	x_1^*	0.940108
WMT	x_2^*	0.919993
IBM	x_3^*	1.471402
MU	x_4^*	0.000000
BA	x_5^*	0.000000
AXP	x_6^*	0.025225
META	x_7^*	3.865146
PEP	x_8^*	1.277443
MCD	x_9^*	0.000000
NFLX	x_{10}^*	0.061941
ADI	x_{11}^*	0.000000
BAC	x_{12}^*	0.297218
PFE	x_{13}^*	0.000000
KO	x_{14}^*	0.306777
MA	x_{15}^*	0.065135
TSLA	x_{16}^*	2.552543
HPQ	x_{17}^*	0.272763
STT	x_{18}^*	0.000000
CMG	x_{19}^*	0.000000
PSX	x_{20}^*	1.303415

Tabuľka 4.2: Výsledné rozloženie portfólia pre $\omega = 0.02$.

Môžme si všimnúť, že software vyhodnotil, že niektoré akcie, ktoré sme uvažovali, nie sú pri replikácií indexu potrebné. Ďalej sa pozrieme na grafické znázornenie vývoju indexu S&P500 a nášho replikačného portfólia v období od 08.06.2020 do 12.12.2022 zobrazeného na obrázku 4.2.



Obr. 4.2: Vývoj indexu S&P 500 a replikačného portfólia v období $in - sample$.

4.3 Vyhodnotenie optimalizačnej úlohy

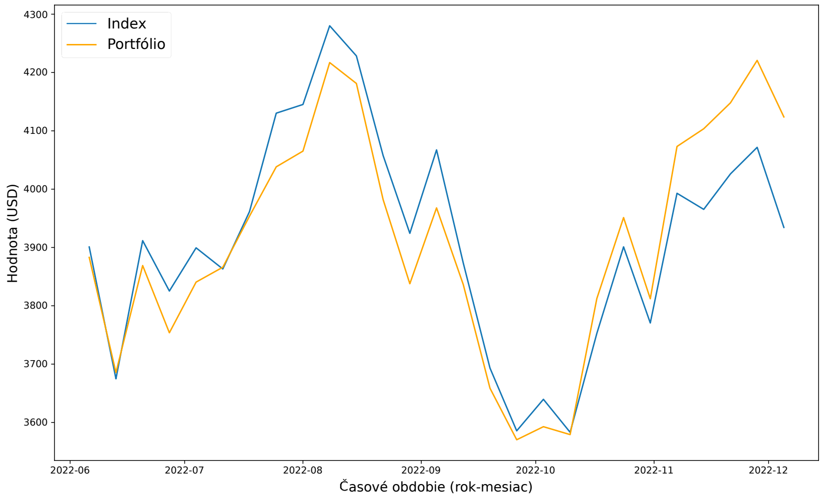
Dôležitou bude pre nás, práve fáza $out-of-sample$, na základe ktorej zistíme nakoľko presné boli naše výpočty v predošlej fáze. Pri výpočte budeme uvažovať portfólio ktorého zloženie pozostáva práve z výsledných hodnôt pre $x_1^*, \ldots, x_{20}^*$ uvedených v tabuľke 4.2. Následne sa pozrieme na vývoj cien jednotlivých akcií v časovom období od 06.06.2022 do 12.12.2022 predstavujúce 27 týždňov a hodnotu, ktorú nadobúda účelová funkcia v tejto fáze. Dosadením optimálnych hodnôt do pôvodnej optimalizačnej úlohy, pre hodnotu účelovej funkcie v tomto období dostávame dostávame

$$g(\boldsymbol{x}) = \frac{1}{27}\sum_{t=1}^{27}\eta_{t,0} = 0.0160902.$$

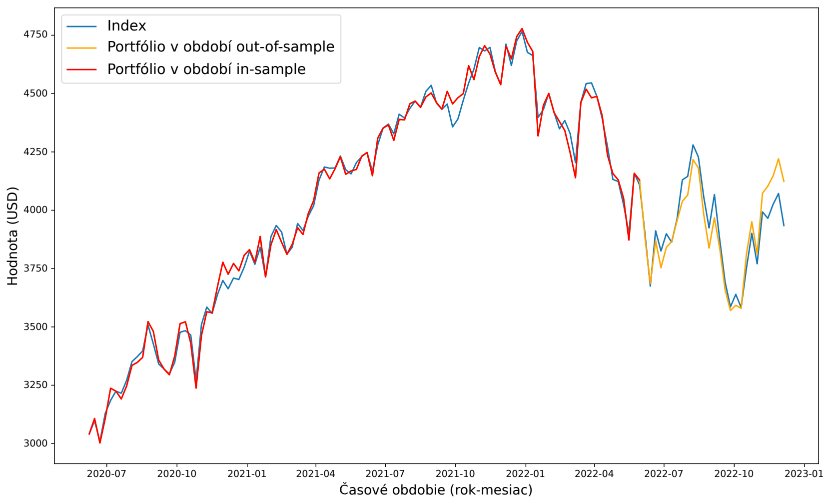
Mieru rizika $CVaR$, ktorú portfólio dosahuje vyčíslime na základe vety 2 pomocou rovnosti

$$\phi_\alpha(\boldsymbol{x}) = \frac{1}{1-\alpha}\Bigg[\Bigg(\sum_{k=1}^{25} p_k - \alpha\Bigg) z_{k_\alpha} + \sum_{k=26}^{27} p_k z_k\Bigg] = 0.02393,$$

kde $p_k = \frac{1}{27}$ pre $k = 1, \ldots, 27$ a z_{k_α} zodpovedá miere rizika VaR v období $out - of - sample$.



Obr. 4.3: Vývoj indexu S&P 500 a replikačného portfólia v období $out - of - sample$.



Obr. 4.4: Vývoj indexu S&P 500 a replikačného portfólia v období od 08.06.2020 do 12.12.2022.

Z obrázka 4.3 môžme vidieť pomerne dobrú zhodu vývoja replikačného portfólia vzhľadom k indexu. Dokonca môžme pozorovať, že portfólio približne od polky desiateho mesiaca roku 2022 nadobúdalo väčšiu hodnotu než samostatný index. Pre lepšiu ukážku sa môžme pozrieť na obrázok 4.4, ktorý zobrazuje celé časové obdobie od 08.06.2020 do 12.12.2022.

Čo sa týka porovnania hodnôt účelových funkcií ich rozdiel pre jednotlivé fázy zodpovedá rozdielu

$$0.0160902 - 0.0063082 = 0.009782.$$

To znamená, že v období $out - of - sample$ bol celkový súčet odchýliek replikačného portfólio od indexu o 0.009782 väčší než v období $in - sample$. A pre hodnoty $CVaR$ platí

$$0.02393 - \omega = 0.02393 - 0.02 = 0.00393,$$

z kadiaľ dostávame, že podmienená hodnota v riziku bola o 0.00393 vyššia.

Po prevedení všetkých výpočtov a nahliadnutím na obrázky sme sa dostali k záveru, že na základe historických dát v rozsahu 104 týždňov s použitím 20 akcií je možné v dostatočne dobrej miere replikovať index S&P 500. Tento záver môžeme usúdiť na základe výsledných hodnôt, ktoré portfólio dosahovalo následujúcich 27 týždňov v období $out - of - sample$, kde dokonca miestami malo väčšiu hodnotu ako samostatný index.

Je dôležité však poznamenať, že ak by sme chceli predikovať vývoj portfólia na dlhšie časové obdobie, bolo by pravdepodobne potrebné na dosiahnutie podobne dobrých výsledkov použiť historické dát vo väčšom rozsahu.

Záver

V tejto práci sme sa zoznámili s viacerými pojmami, ktoré sú bežne používané v oblasti financií ako napríklad pojmy portfólio a riziko. Definovali sme ich matematické formulácie, kde sme sa následne zamerali na jednotlivé metódy merania rizika, ktorými boli VaR a $CVaR$. Na základe vlastností týchto metód sme odvodili funkciu, ktorá nám umožnila efektívne vypočítať riziká súvisiace s portfóliom. Hlavným cieľom tejto práce bolo skonštruovať portfólio, ktoré bude schopné replikovať zvolený index S&P 500.

Na riešenie tohto problému sme využili poznatky z oblasti optimalizácie a zostavili sme lineárny program s účelovou funkciou minimalizácie rizika. Konkrétne sme sa snažili minimalizovať súčet absolútnych hodnôt odchýliek replikačného portfólia od indexu.

V procese konštrukcie programu sme aplikovali aj odvodenú funkciu pre jednotlivé riziká ako jednu z podmienok minimalizácie. Táto funkcia bola obmedzená hodnotou, ktorá predstavovala maximálne riziko, ktoré sme ochotní tolerovať. Výpočet sme prvotne previedli na historických dátach a následne sme sledovali vývoj indexu a replikačného portfólia v následujúcich časových obdobiach.

V praktickej aplikácií sme teda ukázali akou formou vieme narábať s rizikom a využiť ho pri riešení optimalizačnej úlohy v oblasti financií. Výsledky tejto práce môžu byť potenciálne užitočné pre investičné spoločnosti a jednotlivých investorov, ktorí sa zaujímajú o efektívne využitie ich financií.

Literatúra

Dupačová, J. (2002). *Stochastic modeling in economics and finance*. Applied optimization ; 75. Kluwer Academic Publishers, Dordrecht. ISBN 1402008406.

Rockafellar, R. T. a Uryasev, S. (2000). Optimization of conditional value-at risk. *Journal of Risk*, **3**, 21–41.

Rockafellar, R. a Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, **26**(7), 1443–1471.

Zoznam obrázkov

Zoznam tabuliek

Príloha

Kód použitý na riešenie optimalizačnej úlohy v jazyku Python pomocou balíčka PuLP.

```
import yfinance as yf
import numpy as np
import pandas as pd
from pulp import *

#PERIÓDA V KTOREJ PRACUEJME S TÝŽDENNÝMI CENAMI
start_day = '2020-6-8'#pondelok
end_day = "2022-6-6"    #pondelok
casovy_interval ="1wk" #týždňový interval


#CENY INDEXU S&P
cena_indexu = pd.DataFrame(yf.download("^GSPC",start_day,end_day,
                            interval = casovy_interval)['Adj Close'])

#ukážka prvých 5 cien
print(cena_indexu.head())

#ZISKANIE DAT CENY AKCII

tickers = ['AAPL', 'WMT', 'IBM', 'MU', 'BA', 'AXP',"META","PEP","MCD",
           "NFLX","ADI","BAC","PFE","KO","MA","TSLA","HPQ","STT","CMG","PSX"]

# Stihanutie pondelkových cien akcií upravených o trhové parametre
ceny_akcii = pd.DataFrame(yf.download(tickers,start_day,end_day,
                            interval = casovy_interval)['Adj Close'])

#ukážka prvých 5 riadkov
print(ceny_akcii.head())

#LINEARNY PROGRAM

#premenné
T= ceny_akcii.shape[0]  #čas sa rovná počtu riadkov
v= 1000 #počiatočná investícia
n= ceny_akcii.shape[1] #počet akcií  pre indexy j= 1,....,n
```

```
alpha = 0.95 #hladina spoľahlivosti
omega = 0.02 #omedzenie pre hodnotu CVaR
theta = v/ cena_indexu.iloc[0,0] #počet jednotiek indexu
pocet_x = np.arange(0,n).tolist()
casy_t =np.arange(0,T).tolist()

#MODEL

model = LpProblem(name="Optimalizacia_Portfolia", sense=LpMinimize)

#rozhodovacie premenné

x = LpVariable.dicts("x", list(range(n)), lowBound=0)
zeta = LpVariable("zeta",lowBound=0)
u = LpVariable.dicts("u", list(range(T)), lowBound=0)
eta =LpVariable.dicts("eta", list(range(T)), lowBound=0)


#účelová funckia
model += lpSum([eta[t] for t in range(T)]) / T

#obmezenia
model += (lpSum([ceny_akcii.iloc[0,j] * x[j] for j in range(n)]) == v)
model += (zeta + 1/((1- alpha)*T)* lpSum([u[t] for t in range(T)]) <= omega)
for t in range(T):
    model += (((theta * cena_indexu.iloc[t,0] - lpSum([ceny_akcii.iloc[t,j] *
                x[j] for j in range(n)])) / (theta * cena_indexu.iloc[t,0])) <= eta[t])
    model += (((-1) * ((theta * cena_indexu.iloc[t,0] - lpSum([ceny_akcii.iloc[t,j] *
                x[j] for j in range(n)]))/ (theta * cena_indexu.iloc[t,0] ))) <= eta[t])
    model += ((((theta * cena_indexu.iloc[t,0] - lpSum([ceny_akcii.iloc[t,j] *
                x[j] for j in range(n)]))/ (theta * cena_indexu.iloc[t,0]) ) - zeta) <= u[t])



model.solve()

print("Minimálna hodnota:", value(model.objective)) #Hodnota účelovej funkcie

x_values = []
for v in model.variables(): #Optimálne hodnoty
    if v.name.startswith("x"):
        x_values.append((v.name, v.varValue))

x_values_sorted = sorted(x_values, key=lambda x: int(x[0].split('_')[1]))
```

```
print("tabuľka hodnôt ")
tabulka = pd.DataFrame(x_values_sorted, columns=["Variable", "Value"],index= tickers)
print(tabulka) # v práci značíme premenné (x^*)_1,...,(x^*)_20

#################################### OUT - OF - SAMPLE ################################
#out of sample data
start_day2= end_day #pre čisto out of sample
end_day2 = "2022-12-12" #pondelok
cena_indexuofs = pd.DataFrame(yf.download("^GSPC",start_day,end_day2,
                    interval = casovy_interval)['Adj Close']) #celé časové obdobie
ceny_akciiofs = pd.DataFrame(yf.download(tickers,start_day,end_day2,
                    interval = casovy_interval)['Adj Close'])#celé časové obdobie

cena_indexuofs2 = pd.DataFrame(yf.download("^GSPC",start_day2,end_day2,
                    interval = casovy_interval)['Adj Close']) #iba ofs
ceny_akciiofs2 = pd.DataFrame(yf.download(tickers,start_day2,end_day2,
                    interval = casovy_interval)['Adj Close']) #iba ofs
Tofs = ceny_akciiofs2.shape[0] #časové obdobie out - of - sample
hodnoty_x =(tabulka.iloc[:,1])


#Hodnoty funkcií  out of sample
hodnoty_eta_ofs = [] #hodnoty odýchliek
for i in range(Tofs):
    etat = (theta * cena_indexuofs2.iloc[i,0] - np.matmul(np.array(ceny_akciiofs2.iloc[i,:]),
            np.array(hodnoty_x)))/ (theta * cena_indexuofs2.iloc[i,0])
    hodnoty_eta_ofs.append(etat)

var_ofs= pd.DataFrame(hodnoty_eta_ofs).quantile(alpha)
print("VaR Out of sample ", var_ofs)


hodnoty_z =sorted(hodnoty_eta_ofs)
hodnoty_z_nad_kvantilom = []
for i in hodnoty_z:
   if i > float(var_ofs):
      hodnoty_z_nad_kvantilom.append(i)

print("Účelová funkcia out of sample")
print((1/Tofs)* sum(np.abs(hodnoty_eta_ofs)))
print("CVaR out of sample podla vety 2 ")
print(1/(1-alpha)*(((25/27)- alpha)*var_ofs +
                   (1/27)*(sum(hodnoty_z_nad_kvantilom))))
```

Výstup kódu.

```
[*********************100%***********************]  1 of 1 completed
              Adj Close
Date
2020-06-08  3041.310059
2020-06-15  3097.739990
2020-06-22  3009.050049
2020-06-29  3130.010010
2020-07-06  3185.040039
[*********************100%***********************]  20 of 20 completed
                 AAPL         ADI        AXP  ...        STT        TSLA         WMT
Date                                          ...
2020-06-08  83.273621  112.690491  97.660835  ...  57.603680   62.352001  112.760124
2020-06-15  85.957664  114.374985  96.950096  ...  59.684418   66.726669  114.780876
2020-06-22  86.918701  112.176575  89.727341  ...  55.240028   63.982666  113.315582
2020-06-29  89.494568  115.383781  90.601372  ...  57.612804   80.577332  114.167946
2020-07-06  94.304688  118.486313  89.951149  ...  59.014786  102.976669  125.152817

[5 rows x 20 columns]
Welcome to the CBC MILP Solver
Version: 2.10.3
Build Date: Dec 15 2019

command line -
↪   /Users/andreapolakovicova/venv/lib/python3.8/site-packages/pulp/solverdir/cbc/osx/64/cbc
↪   /var/folders/1f/y5fj1q954p3627tp53_shsw00000gn/T/4d354254868d45349399cde2b3a8835e-pulp.mps
↪   timeMode elapsed branch printingOptions all solution
↪   /var/folders/1f/y5fj1q954p3627tp53_shsw00000gn/T/4d354254868d45349399cde2b3a8835e-pulp.sol
↪   (default strategy 1)
At line 2 NAME          MODEL
At line 3 ROWS
At line 319 COLUMNS
At line 7205 RHS
At line 7520 BOUNDS
At line 7521 ENDATA
Problem MODEL has 314 rows, 229 columns and 6781 elements
Coin0008I MODEL read with 0 errors
Option for timeMode changed from cpu to elapsed
Presolve 314 (0) rows, 229 (0) columns and 6781 (0) elements
Perturbing problem by 0.001% of 0.012714866 - largest nonzero change 5.0546408e-06 ( 0.05187479%)
↪   - largest zero change 5.0376604e-06
0  Obj 0 Primal inf 188.92781 (209)
31  Obj 0.00090556942 Primal inf 40.944698 (126)
78  Obj 0.0046607834 Primal inf 8.6433813 (118)
125  Obj 0.0059080194 Primal inf 0.83329501 (51)
161  Obj 0.006314715
Optimal - objective value 0.0063082578
Optimal objective 0.006308257763 - 161 iterations time 0.012
Option for printingOptions changed from normal to all
Total time (CPU seconds):       0.01   (Wallclock seconds):       0.02

Minimálna hodnota: 0.006308257776355769
tabuľka hodnôt
     Variable     Value
AAPL      x_0  0.940105
WMT       x_1  0.919989
IBM       x_2  1.476853
MU        x_3  0.000000
BA        x_4  0.000000
AXP       x_5  0.025225
META      x_6  3.865161
PEP       x_7  1.277447
MCD       x_8  0.000000
NFLX      x_9  0.062037
ADI      x_10  0.000000
BAC      x_11  0.297218
PFE      x_12  0.000000
KO       x_13  0.306778
MA       x_14  0.065141
TSLA     x_15  2.552534
HPQ      x_16  0.272761
STT      x_17  0.000000
```

```
CMG      x_18  0.000000
PSX      x_19  1.303414
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  20 of 20 completed
[*********************100%***********************]  1 of 1 completed
[*********************100%***********************]  20 of 20 completed
VaR Out of sample  0    0.022205
Name: 0.95, dtype: float64
Účelová funkcia out of sample
0.016090249438966323
CVaR out of sample podla vety 2
0    0.02393
Name: 0.95, dtype: float64

Process finished with exit code 0
```