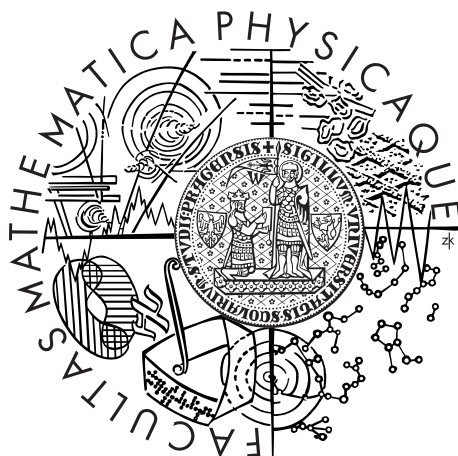


Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## BAKALÁŘSKÁ PRÁCE



Hana Koláčková

## Generování scénářů při požadavku na shodu momentů

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Jitka Dupačová, DrSc.

Studijní program: Matematika

Studijní obor: Finanční matematika

Praha 2015

Ráda bych poděkovala své vedoucí prof. RNDr. Jitce Dupačové, DrSc. a konzultantovi RNDr. Václavu Kozmíkovi, Ph.D. za odbornou konzultaci a poskytnutí cenných rad v průběhu psaní 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                      Praze                      dne    19. 5. 2015

Název práce: Generování scénářů při požadavku na shodu momentů

Autor: Hana Koláčková

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Jitka Dupačová, DrSc.

Abstrakt:

V práci jsou uvedeny čtyři způsoby generování scénářů tak, aby výsledné diskrétní rozdělení pravděpodobnosti replikovalo předepsané hodnoty momentů. Prvním z nich je heuristický algoritmus, druhým ze způsobů je symetrické rozložení okolo střední hodnoty, další je systém nelineárních rovnic a poslední způsob je řešení pomocí cílového programování. V další části je přiblížen charakter této úlohy cílového programování a jsou uvedeny konkrétní možnosti specifikací parametrů úlohy s následným ovlivněním náročnosti řešení. V poslední části práce jsou porovnány výsledky několika možných způsobů řešení vybraných typů úlohy.

Klíčová slova: Cílové programování, scénáře, shoda momentů, heuristický algoritmus, rovnoměrné rozložení okolo střední hodnoty

Title: Scenario generation by the moment fitting method

Author: Hana Koláčková

Department: Department of Probability and Mathematical Statistics

Supervisor: prof. RNDr. Jitka Dupačová, DrSc.

Abstract:

The thesis presents four methods for scenario generating leading to the resulting discrete probability distribution that replicates given values of the moments. The first method uses heuristic algorithm, the second method generates by symmetrically distributing values around the mean value, the third one is based on solving the system of nonlinear equations and finally the last method is based on goal programming. Next section describes the nature of problems solved by the goal programming. It also details possible ways of parameter specification to allow control of the computational complexity. In the last part of the thesis the results of several suitable methods for chosen types of problem are compared.

Keywords: Goal programming, heuristic algorithm, scenarios, fitting moments

# Obsah

<b>1</b>	<b>Nastolení problému</b>	<b>3</b>
1.1	Úvod . . . . .	3
1.2	Základní definice, věty a vlastnosti . . . . .	4
1.2.1	Náhodné veličiny a vektory . . . . .	4
1.2.2	Matice . . . . .	5
1.2.3	Konvexní funkce . . . . .	6
1.3	Problém . . . . .	8
<b>2</b>	<b>Metody řešení</b>	<b>10</b>
2.1	Heuristický algoritmus . . . . .	10
2.2	Pravidelné rozložení scénářů . . . . .	11
2.2.1	Shoda prvních momentů . . . . .	11
2.2.2	Shoda prvních momentů a rozptylů . . . . .	12
2.3	Soustava nelineárních rovnic . . . . .	13
2.4	Řešení pomocí cílového programování . . . . .	13
<b>3</b>	<b>Charakter úlohy (2.9)</b>	<b>15</b>
3.1	Speciální volba rozměrů úlohy . . . . .	16
3.2	Volba metriky . . . . .	17
3.2.1	$L_1$ norma . . . . .	17
3.2.2	$L_2$ norma . . . . .	20
<b>4</b>	<b>Praktická část</b>	<b>23</b>
4.1	Data . . . . .	23
4.2	Shoda prvních momentů . . . . .	24
4.2.1	Lineární programování . . . . .	24
4.2.2	Výběr z reálných dat . . . . .	26
4.2.3	Scénáře nalezené pro shodu prvních momentů . . . . .	27
4.3	Shoda druhých momentů . . . . .	29
4.3.1	Nelineární programování . . . . .	29
4.3.2	Výběr z reálných dat . . . . .	30
4.3.3	Scénáře nalezené pro shodu druhých momentů . . . . .	31
4.4	Shoda středních hodnot a rozptylů . . . . .	33
4.4.1	Pravidelné rozložení scénářů . . . . .	33
4.4.2	Nelineární programování . . . . .	34
4.4.3	Výběr z reálných dat . . . . .	35
4.4.4	Scénáře nalezené pro shodu prvních momentů a rozptylů . . . . .	36
4.5	Shoda středních hodnot, rozptylů a korelací . . . . .	37
4.6	Časová náročnost výpočtů . . . . .	39

4.7 Závěr . . . . .	40
<b>Literatura</b>	<b>41</b>
<b>Seznam obrázků</b>	<b>42</b>
<b>Seznam tabulek</b>	<b>43</b>

# Kapitola 1

## Nastolení problému

### 1.1 Úvod

V dnešní době je investování velmi rozšířenou aktivitou jak mezi běžnými lidmi tak mezi odborníky, kteří se jím přímo živí. Hlavním problémem této oblasti je ale rozhodnout se, kam investovat a kam naopak neinvestovat. Toto rozhodnutí je náročné především kvůli nejistotě ohledně budoucího vývoje komodity, do které chceme investovat. Navíc bývá nutné se rozhodnout nejen na jedno ale i na více časových období, čímž se rozhodnutí stává ještě komplikovanějším. Typickými příklady důvodů znemožňujících jednoduché rozhodování jsou proměnlivé úrokové míry, inflace nebo kurzy měn. Proto byly vytvořeny stochastické modely, které mají za cíl vysvětlit nebo popsat náhodné parametry používané při vstupu do modelů, které tato rozhodnutí ovlivňují.

Velmi důležitou složkou práce s takovými modely je reprezentovat parametry ve formě vhodné pro výpočty. Představme si například, že máme velký soubor dat ze stejného rozdělení, ať už je spojitě nebo diskrétní. Pro výpočet střední hodnoty nebo i jakéhokoliv dalšího statistického parametru potřebujeme vypočítat mnohorozměrný integrál v případě spojitěho rozdělení a několiknásobnou sumu v případě diskrétního rozdělení a ty ani nemusejí být explicitně definovány. Takové výpočty jsou náročné a z numerického hlediska neschůdné.

Z tohoto důvodu bychom rádi nahradili takto výpočetně náročné rozdělení nějakým jiným, pro výpočty vhodnějším rozdělením. To znamená rozdělením, které bude reprezentováno méně hodnotami, aby byly výpočty jednodušší, ale zároveň takovým, které bude mít stejné alespoň některé statistické vlastnosti jako například marginální momenty nebo kovariance. Jednotlivé body tohoto nového rozdělení nazveme scénáře, přičemž každému přísluší jeho pravděpodobnost.

Na jednotlivé scénáře můžeme pohlížet jako na jednotlivé atomy diskrétního rozdělení. Jeden scénář je jedním atomem, bodem daného diskrétního rozdělení. Jejich počet sice není v podstatě ničím omezen, ale my se snažíme o co nejmenší počet, aby byly výpočty co nejjednodušší a nejrychlejší, ale zároveň jich potřebujeme dostatečné množství, aby dobře reprezentovaly dané rozdělení. Pojmem „dobře“ zde rozumíme tak, aby zůstaly zachovány námi zvolené statistické vlastnosti jako například střední hodnota, rozptyl nebo i vyšší momenty.

## 1.2 Základní definice, věty a vlastnosti

### 1.2.1 Náhodné veličiny a vektory

**Definice 1.** [2, Kapitola 2.2] Nechť  $\Omega$  je nějaká množina a  $\mathcal{A}$  je  $\sigma$ -algebra jejích podmnožin. Funkci  $P : \mathcal{A} \rightarrow \langle 0, 1 \rangle$  nazveme pravděpodobností, právě když splňuje následující podmínky:

- (a)  $P(A) \geq 0, P(\Omega) = 1, \forall A \in \mathcal{A}$ ;
- (b)  $A_1, A_2, A_3, \dots \in \mathcal{A}$  a  $A_i \cap A_j = \emptyset \forall i \neq j \Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ .

**Definice 2.** [1, Kapitola 1.1] Rozdělením náhodné veličiny  $X : (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$ , kde  $\mathcal{X}$  je nějaká množina a  $\mathcal{B}$  je nějaká  $\sigma$ -algebra na  $\mathcal{X}$ , rozumíme indukovanou pravděpodobnostní míru  $P_X$  na  $(\mathcal{X}, \mathcal{B})$  definovanou vztahem

$$P_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\}), B \in \mathcal{B}.$$

**Definice 3.** [1, Kapitola 1.1] Střední hodnotou (reálné) náhodné veličiny  $X$  rozumíme číslo  $E X$  dané výrazem

$$E X = \int_{\Omega} X(\omega) dP(\omega),$$

pokud integrál na pravé straně existuje. Hodnota  $\mu_k = E X^k$  se nazývá  $k$ -tý moment náhodné veličiny  $X$ . Hodnota  $\mu_k^* = E(X - E X)^k$  se nazývá  $k$ -tý centrální moment náhodné veličiny  $X$ . Druhý centrální moment náhodné veličiny  $X$  se nazývá rozptyl a značí se  $\text{var } X$ .

**Definice 4.** [1, Kapitola 2.1] Nechť  $X, Y$  jsou náhodné veličiny s konečnými druhými momenty. Kovarianci veličin  $X$  a  $Y$  definujeme jako

$$\text{cov}(X, Y) = E(X - E X)(Y - E Y).$$

**Definice 5.** [1, Kapitola 2.6] Nechť  $X, Y$  jsou náhodné veličiny s konečnými druhými momenty a s kladnými rozptily. Korelační koeficient veličin  $X$  a  $Y$  je

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var } X \text{var } Y}}.$$

**Definice 6.** [1, Kapitola 2.1] Nechť náhodné veličiny  $X_1, \dots, X_n$  jsou definovány na témž pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$ . Pak vektor  $\mathbf{X} = (X_1, \dots, X_n)^\top$  nazveme náhodným vektorem.

**Definice 7.** [1, Kapitola 2.1] Mějme pravděpodobnostní prostor  $(\Omega, \mathcal{A}, P)$  a náhodný vektor  $\mathbf{X} : \Omega \rightarrow \mathbb{R}^n$  v  $(\Omega, \mathcal{A}, P)$ . Pak se pravděpodobnostní míra  $P_{\mathbf{X}} : \mathcal{B}^n \rightarrow \mathbb{R}$  na  $\sigma$ -algebře  $n$ -rozměrných Borelovských množin  $\mathcal{B}^n$  definovaná

$$P_{\mathbf{X}}(B) = P(\{\mathbf{X} \in B\}), \text{ pro každou } B \in \mathcal{B}^n,$$

nazývá rozdělení pravděpodobnosti náhodného vektoru  $\mathbf{X}$ .

- Rozdělení celého náhodného vektoru  $\mathbf{X} = (X_1, \dots, X_n)^\top$  se nazývá sdružené rozdělení.



- Necht  $k_1, \dots, k_r$  jsou různá celá čísla,  $1 \leq k_i \leq n$  pro  $i = 1, 2, \dots, r$ , přičemž  $1 \leq r < n$ . Rozdělení náhodných veličin  $(X_{k_1}, \dots, X_{k_r})^\top$  se pak nazývá marginální.

**Definice 8.** [1, Kapitola 2.1] Mějme náhodný vektor  $\mathbf{X} = (X_1, \dots, X_n)^\top$  a necht existují střední hodnoty  $E X_1, E X_2, \dots, E X_n$ . Pak

$$E \mathbf{X} = (E X_1, \dots, E X_n)^\top$$

nazveme střední hodnota náhodného vektoru  $\mathbf{X} = (X_1, \dots, X_n)^\top$ . Hodnota  $\mu_k = E \mathbf{X}^k = (E(X_1)^k, \dots, E(X_n)^k)^\top$  se nazývá  $k$ -tý moment náhodného vektoru  $\mathbf{X}$ .

**Definice 9.** [1, Kapitola 2.1] Necht náhodné vektory  $\mathbf{X} = (X_1, \dots, X_n)^\top$  a  $\mathbf{Y} = (Y_1, \dots, Y_m)^\top$  mají konečné druhé momenty. Kovarianci náhodného vektoru  $\mathbf{X}$  náhodných veličin  $X_i$  a  $X_j$  ( $1 \leq i, j \leq n$ ) definujeme vztahem

$$\text{cov}(X_i, X_j) = E(X_i - E X_i)(X_j - E X_j).$$

Kovarianční matice  $\text{cov}(\mathbf{X}, \mathbf{Y})$  vektorů  $\mathbf{X}$  a  $\mathbf{Y}$  je definovaná vzorcem

$$\text{cov}(\mathbf{X}, \mathbf{Y}) = E(\mathbf{X} - E \mathbf{X})(\mathbf{Y} - E \mathbf{Y})^\top$$

**Definice 10.** [1, Kapitola 2.6] Necht  $\mathbf{X} = (X_1, \dots, X_n)^\top$  a  $\mathbf{Y} = (Y_1, \dots, Y_m)^\top$  jsou dva náhodné vektory se složkami, jež mají konečné a kladné rozptyly. Korelační maticí  $\rho(\mathbf{X}, \mathbf{Y})$  vektorů  $\mathbf{X}$  a  $\mathbf{Y}$  rozumíme matici typu  $n \times m$  se složkami  $\rho(X_i, Y_j)$  na místě  $(i, j)$ .

## 1.2.2 Matice

**Definice 11.** [4, Kapitola 1.1] Čtvercová matice  $A$  stupně  $n$  se nazývá symetrická, jestliže  $A = A^\top$ , tj.  $a_{ij} = a_{ji}$ ,  $i, j = 1, 2, \dots, n$ .

**Definice 12.** [4, Kapitola 1.1] Čtvercovou matici  $A$  nazveme regulární, pokud existuje matice  $B$  taková, že  $AB = BA = I$ , kde  $I$  je jednotková matice. Matici  $B$  nazveme inverzní maticí k matici  $A$ . Čtvercovou matici, která není regulární, nazveme singulární maticí.

**Definice 13.** [10, Definice 9.10] Je-li  $A$  symetrická matice typu  $n \times n$ , pak funkci  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , definovanou předpisem

$$\phi(h) = (Ah, h) = \sum_{i,j=1}^n a_{ij} h_i h_j$$

nazýváme kvadratickou formou danou maticí  $A$ . Tato kvadratická forma se nazývá

- pozitivně definitní, je-li  $\phi(h) > 0 \forall h \in \mathbb{R}^n, h \neq 0$ ,
- pozitivně semidefinitní, je-li  $\phi(h) \geq 0 \forall h \in \mathbb{R}^n$ ,
- indefinitní, nabývá-li jak kladných tak záporných hodnot.

**Věta 1.** [4, Věta 4.8] Necht je  $A$  pozitivně definitní symetrická reálná čtvercová matice. Pak existuje jednoznačný rozklad  $A = LL^\top$ , kde  $L$  je dolní trojúhelníková matice s kladnými prvky na diagonále. Tento jednoznačně určený rozklad nazveme Choleského rozklad.

### 1.2.3 Konvexní funkce

**Definice 14.** [10, Definice 8.22] Necht  $X$  je libovolná neprázdná množina a  $d$  je funkce přiřazující každé dvojici  $(x,y)$  z  $X$  nezáporné reálné číslo  $d(x,y)$  s vlastnostmi:

- a) pro  $x,y \in X$  je  $d(x,y) = 0$  právě když  $x = y$ ;
- b)  $d(x,y) = d(y,x)$  pro každé  $x,y \in X$ ;
- c)  $d(x,y) \leq d(x,z) + d(z,y)$  pro každé  $x,y,z \in X$ .

Funkce  $d$  se pak nazývá metrika na  $X$  a dvojice  $(X,d)$  se nazývá metrický prostor.

**Definice 15.** [12, Kapitola 1.3] Necht  $\mathbf{x} = (x_1, \dots, x_n)^\top$  a  $p > 1$  je reálné číslo, pak  $p$ -normu vektoru  $\mathbf{x}$  definujeme vztahem

$$\|\mathbf{x}\|_p := \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

**Definice 16.** [13, 10.1] Necht  $p < \infty$ . Množinu všech  $\mu$ -měřitelných funkcí na  $X$  takových, že

$$\int_X |f|^p d\mu < \infty,$$

označme  $\mathcal{L}^p = \mathcal{L}^p(X, \varphi, \mu)$ , kde  $\varphi$  je  $\sigma$ -algebra. Pak hodnota

$$\|f\|_p := \left( \int_X |f|^p d\mu \right)^{\frac{1}{p}}$$

se nazývá  $L_p$ -norma funkce  $f \in \mathcal{L}^p$ .

**Definice 17.** [6, Definice 2.2] Množina  $D \subset \mathbb{R}^n$  se nazývá konvexní, jestliže s každými dvěma body obsahuje také všechny jejich konvexní lineární kombinace, tj.  $\forall \mathbf{x}, \mathbf{y} \in D$  a  $\lambda \in (0,1)$  je také  $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in D$ .

**Definice 18.** [6, Definice 2.11] Necht  $D \subset \mathbb{R}^n$  je konvexní množina, funkce  $f : D \rightarrow \mathbb{R}$  se nazývá konvexní na  $D$ , jestliže  $\forall \mathbf{x}, \mathbf{y} \in D$  a  $\forall \lambda \in [0,1]$  platí:

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

**Věta 2.** Necht  $D \subset \mathbb{R}^n$  je konvexní množina a  $f_1 : D \rightarrow \mathbb{R}, f_2 : D \rightarrow \mathbb{R}, \dots, f_k : D \rightarrow \mathbb{R}$  jsou konvexní funkce. Pak pro každé váhy  $a_1 \geq 0, a_2 \geq 0, \dots, a_k \geq 0$  je  $\sum_{i=1}^k a_i f_i$  opět konvexní funkcí.

*Důkaz.* Důkaz provedeme sporem. Necht tedy  $\sum_{i=1}^k a_i f_i$  není konvexní, pak  $\exists \mathbf{x}, \mathbf{y} \in D$  a  $\alpha \in (0,1)$  tak, že

$$\alpha \sum_{i=1}^k a_i f_i(\mathbf{x}) + (1 - \alpha) \sum_{i=1}^k a_i f_i(\mathbf{y}) < \sum_{i=1}^k a_i f_i(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}).$$

To úpravou převedeme na tvar

$$\sum_{i=1}^k a_i (\alpha f_i(\mathbf{x}) + (1 - \alpha) f_i(\mathbf{y})) < \sum_{i=1}^k a_i f_i(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}).$$

Pak alespoň pro jedno  $i \in \{1, \dots, k\}$  platí

$$a_i(\alpha f_i(\mathbf{x}) + (1 - \alpha)f_i(\mathbf{y})) < a_i f_i(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}).$$

Vydělením  $a_i > 0$  získáme

$$\alpha f_i(\mathbf{x}) + (1 - \alpha)f_i(\mathbf{y}) < f_i(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}).$$

To je spor s předpokladem konvexity všech funkcí  $f_i$ , tedy  $\sum_{i=1}^k a_i f_i$  je konvexní funkce na  $D$ .  $\square$

**Věta 3.** [6, Lemma 2.38] Necht' je  $D \subset \mathbb{R}^n$  konvexní množina,  $h : D \rightarrow \mathbb{R}$  je afinně lineární funkce a  $g : \mathbb{R} \rightarrow \mathbb{R}$  je konvexní funkce. Pak  $f : D \rightarrow \mathbb{R} : \mathbf{x} \rightarrow g(h(\mathbf{x}))$  je konvexní funkce.

**Věta 4.** [6, Lemma 2.39] Necht' je  $D \subset \mathbb{R}^n$  konvexní množina,  $h : D \rightarrow \mathbb{R}$  je konvexní funkce a  $g : \mathbb{R} \rightarrow \mathbb{R}$  je neklesající konvexní funkce. Pak  $f : D \rightarrow \mathbb{R} : \mathbf{x} \rightarrow g(h(\mathbf{x}))$  je konvexní funkce.

**Definice 19.** [10, Definice 9.7] Necht' funkce  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  je definovaná na nějakém okolí bodu  $\mathbf{a} \in \mathbb{R}^n$  na množině  $M \in \mathbb{R}^n, \mathbf{a} \in M$ . Řekneme, že funkce  $f$  má v  $\mathbf{a}$ :

- lokální minimum vzhledem k  $M$ , jestliže existuje takové okolí  $U(\mathbf{a})$ , že  $f(\mathbf{a}) \leq f(\mathbf{x})$  pro  $\mathbf{x} \in U(\mathbf{a}) \cap M$ ,
- ostré lokální minimum vzhledem k  $M$ , jestliže existuje takové okolí  $U(\mathbf{a})$ , že  $f(\mathbf{a}) < f(\mathbf{x})$  pro  $\mathbf{x} \in U(\mathbf{a}) \cap M$ ,
- globální minimum vzhledem k  $M$ , jestliže  $f(\mathbf{a}) \leq f(\mathbf{x}) \forall \mathbf{x} \in M, \mathbf{x} \neq \mathbf{a}$ ,
- ostré globální minimum vzhledem k  $M$ , jestliže  $f(\mathbf{a}) < f(\mathbf{x}) \forall \mathbf{x} \in M, \mathbf{x} \neq \mathbf{a}$ .

**Definice 20.** [10, Definice 9.1] Necht' funkce  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  je definovaná na nějakém okolí bodu  $\mathbf{a} \in \mathbb{R}^n, \mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ . Potom parciální derivací funkce  $f$  podle  $i$ -té proměnné v bodě  $\mathbf{a}$  nazýváme vlastní limitu (pokud existuje)

$$\lim_{h \rightarrow 0} = \frac{f(a_1, a_2, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(\mathbf{a})}{h}.$$

Označujeme ji  $f_{x_i}(\mathbf{a})$ .

**Věta 5.** [10, Věta 9.22] Necht' funkce  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  je definovaná na  $\mathcal{U}(\mathbf{a})$ , kde  $\mathbf{a}$  je nějaký bod z  $\mathbb{R}^n$  a  $\mathcal{U}(\mathbf{a})$  je okolí bodu  $\mathbf{a}$ . Má-li  $f$  v bodě  $\mathbf{a}$  lokální extrém a má v něm derivaci  $f_{x_i}(\mathbf{a})$  podle  $i$ -té proměnné, pak je tato derivace rovna nule.

**Věta 6.** Necht'  $f$  je nekonstantní funkce a má ve vnitřním bodě  $c$  svého definičního oboru lokální extrém. Je-li  $f$  v okolí bodu  $c$  konvexní, má v  $c$  minimum.

*Důkaz.* Důkaz provedeme sporem. Necht'  $\forall \mathbf{y} \in \mathcal{U}_{\mathbf{x}} \exists \mathbf{z} \in \mathcal{U}_{\mathbf{x}} : \mathbf{x} = \alpha \mathbf{y} + (1 - \alpha)\mathbf{z}$ , kde  $\mathcal{U}_{\mathbf{x}}$  je okolí bodu  $\mathbf{x}$  a  $\alpha \in (0, 1)$ . Necht' ještě není v  $\mathbf{x}$  lokální minimum (tj. je v něm lokální maximum), pak platí  $\forall \mathbf{y} \in \mathcal{U}_{\mathbf{x}}, f(\mathbf{y}) \leq f(\mathbf{x})$ .  $f$  je konvexní na  $\mathcal{U}_{\mathbf{x}}$ , to znamená, že

$$\alpha f(\mathbf{y}) + (1 - \alpha)f(\mathbf{z}) \geq f(\alpha \mathbf{y} + (1 - \alpha)\mathbf{z}) = f(\mathbf{x})$$

tj. alespoň jedna z  $f(\mathbf{y}), f(\mathbf{z}) > f(\alpha \mathbf{y} + (1 - \alpha)\mathbf{z}) = f(\mathbf{x})$  a to je spor s předpokladem, že v  $\mathbf{x}$  je lokální maximum. Tedy je v něm lokální minimum.  $\square$

**Věta 7.** [6, Lemma 2.33] *Nechť  $\mathcal{X} \subset \mathbb{R}^n$  je otevřená konvexní množina a funkce  $f : \mathcal{X} \rightarrow \mathbb{R}$  má druhé parciální derivace v každém bodě  $\mathcal{X}$ . Pak  $f$  je konvexní funkce tehdy a jen tehdy, když její matice druhých parciálních derivací je pozitivně semidefinitní v každém bodě  $x \in \mathcal{X}$ .*

**Věta 8.** [6, Věta 2.37] *Nechť  $\mathcal{X} \subset \mathbb{R}^n$  konvexní a nechť  $f : \mathcal{X} \rightarrow \mathbb{R}$  je konvexní funkce. Pak každé její lokální minimum je minimem globálním.*

## 1.3 Problém

Předpokládejme, že o původním rozdělení máme úplnou znalost, ač to není úplně realistický předpoklad. Pokud je toto rozdělení diskrétní, můžeme rovnou přejít k výběru vhodných scénářů. V případě, že je toto známé rozdělení spojitě, nahradíme jej rozdělením diskrétním. To je možné provést například simulací z původního rozdělení. Tak budeme opět v situaci, kdy pracujeme s diskrétním rozdělením, které je ale stále nepříjemné pro výpočty, a proto jej chceme nahradit rozdělením s menším počtem scénářů  $\omega^1, \dots, \omega^S, S < \infty$  s příslušnými pravděpodobnostmi  $p^1, \dots, p^S$ .

Další možností je, že nemáme úplnou znalost daného pravděpodobnostního rozdělení. Například známe jen některé momenty. V takovémto případě je nutné vhodně a dobře zvolit přibližné rozdělení při dané informaci.

Otázka možné reprezentace pravděpodobnostního rozdělení pomocí nekonečné sekvence momentů a její aproximace užitím pouze několika z těchto momentů je spojena s momentovým problémem. Navíc je možné dokázat, že vzhledem k  $m$  přípustným hodnotám momentů existuje diskrétní pravděpodobnostní rozdělení s těmito momenty a nosičem majícím nanejvýš  $m + 2$  bodů.

**Věta 9.** [14, Theorém 5.1.4] *Nechť  $\sigma$ -algebra na  $\Omega$  existuje, pravděpodobnostní míra  $P$  je na ní definovaná a  $u_0(z), \dots, u_m(z), z \in \Omega$  jsou měřitelné a integrovatelné funkce vzhledem k  $\sigma$ -algebře a míře  $P$ , pak existuje pravděpodobnostní míra  $P'$  na konečném nosiči  $\Omega'$  taková, že*

$$\mu'_k = \int_{\Omega} u_k(z) dP' = \mu_k, k = 0, \dots, m,$$

kde  $\mu_k = \int_{\Omega} u_k(z) dP, k = 0, \dots, m$  a nosič pravděpodobnostní míry  $P'$  má nejvíce  $m + 2$  bodů.

To tedy znamená, že pro dané hodnoty konkrétních momentů, například středních hodnot, funkce označené jako  $\mu_k = \int_{\Omega} u_k(z) dP(z), k = 1, \dots, m$  existuje nevelký počet scénářů  $\omega^s, s = 1, \dots, S$  a jejich pravděpodobností  $p^s, s = 1, \dots, S : \sum_{s=1}^S p^s = 1$  tak, že hodnoty těchto momentů jsou zachovány, to jest  $\sum_{s=1}^S p^s u_k(\omega^s) = \mu_k, k = 1, \dots, m$ .

Abychom získali tyto scénáře a jejich pravděpodobnosti, potřebujeme najít řešení  $\omega^s$  a  $p^s, s = 1, \dots, S$  systému

$$\sum_{s=1}^S p^s u_k(\omega^s) = \mu_k, k = 1, \dots, m \quad (1.1)$$

rozšířeného o podmínky  $p^s \geq 0 \forall s = 1, \dots, S$  a  $\sum_{s=1}^S p^s = 1$ , kde lze systém rovnic dále rozšířit o další omezení. Obecně je tento problém nelineární.

Předpokládejme, že  $\omega = (\omega_1, \dots, \omega_n)$  je  $n$ -rozměrný náhodný vektor s marginálními rozděleními s prvními  $m$  momenty  $\mu_k(1), \mu_k(2), \dots, \mu_k(n), k = 1, \dots, m$  a kovariancemi  $\rho_{ij}, i \neq j, i = 1, \dots, n-1, j = i+1, \dots, n$  sdruženého rozdělení. Abychom pokryli extrémální případy, přidáme podmínku  $\omega_1^s \geq c_1, \omega_2^s \geq c_2, \dots, \omega_n^s \geq c_n$ , přičemž nemusí nutně platit všechny podmínky.

Nechť máme ještě  $n$ -rozměrné pravděpodobnostní rozdělení, které se shoduje s původním v  $S$  atomech  $\omega^s = (\omega_1^s, \dots, \omega_n^s), s = 1, \dots, S$  s pravděpodobnostmi  $p^s$  splňujícími  $p^s \geq 0 \forall s = 1, \dots, S, \sum_{s=1}^S p^s = 1$  a má předepsané hodnoty momentů tj.  $\sum_{s=1}^S p^s u_k(\omega^s) = \mu_k, k = 1, \dots, m$ . Hledání řešení tohoto systému tedy znamená hledání hodnot vektorových  $n$ -tic  $(\omega_1^s, \dots, \omega_n^s), s = 1, \dots, S$  a skalárů  $p^s, s = 1, \dots, S$  tak, aby

$$\sum_{s=1}^S p^s u_k(\omega_i^s) = \mu_k(i), k = 1, \dots, m, i = 1, \dots, n$$

a případně také

$$\sum_{s=1}^S p^s (\omega_i^s - \mu_1(i)) (\omega_j^s - \mu_1(j)) = \text{cov}_{ij}, i = 1, \dots, n-1, j = i+1, \dots, n,$$

pokud chceme zachovat hodnoty kovariancí, při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2, \dots, \omega_n^s \geq c_n, p^s \geq 0, s = 1, \dots, S, \sum_{s=1}^S p^s = 1$ .

Označme  $(\omega_1^s)^k := u_k(\omega^s)$ . Předchozí rovnosti pak můžeme rozepsat jako

$$\begin{aligned} \sum_{s=1}^S p^s (\omega_1^s)^k &= \mu_k(1), k = 1, \dots, m \\ \sum_{s=1}^S p^s (\omega_2^s)^k &= \mu_k(2), k = 1, \dots, m \\ &\vdots \\ \sum_{s=1}^S p^s (\omega_n^s)^k &= \mu_k(n), k = 1, \dots, m \\ \sum_{s=1}^S p^s (\omega_1^s - \mu_1(1)) (\omega_j^s - \mu_1(j)) &= \text{cov}_{1,j}, j = 2, \dots, n \\ \sum_{s=1}^S p^s (\omega_2^s - \mu_1(2)) (\omega_j^s - \mu_1(j)) &= \text{cov}_{2,j}, j = 3, \dots, n \\ &\vdots \\ \sum_{s=1}^S p^s (\omega_{n-1}^s - \mu_1(n-1)) (\omega_n^s - \mu_1(n)) &= \text{cov}_{n-1,n} \end{aligned} \tag{1.2}$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2, \dots, \omega_n^s \geq c_n, p^s \geq 0, s = 1, \dots, S, \sum_{s=1}^S p^s = 1$ .

# Kapitola 2

## Metody řešení

### 2.1 Heuristický algoritmus

Jako možné řešení problému (1.2) navrhli K. Hoyland, M. Kaut a S. W. Wallace efektivní, rychlý a lehce uživatelný heuristický algoritmus založený na transformacích [8, 9], jehož cílem je vygenerovat scénáře a jejich pravděpodobnosti tak, aby měly požadované první čtyři marginální momenty a korelace.

Algoritmus má dvě základní části, inicializační a transformační. V první části je potřeba specifikovat požadované momenty a korelační matici. Ta musí být symetrická pozitivně semi-definitní s jedničkami na hlavní diagonále a také chceme, aby nebyla singulární. K nesplnění těchto podmínek dochází nejčastěji v případě vnitřní inkonzistence v datech nebo v případě kolinearity náhodných proměnných. Dále je potřeba vygenerovat  $n$  realizací jednorozměrných náhodných veličin, které obvykle mají standardizované normální rozdělení, a z nich vytvořit vícerozměrnou diskrétní náhodnou veličinu kombinováním těchto jednorozměrných realizací následujícím způsobem:

- všechna marginální rozdělení jsou generována se stejným počtem realizací
- pravděpodobnost  $i$ -té realizace je stejná pro všechna marginální rozdělení
- $i$ -tý scénář, tedy  $i$ -tou realizaci sdruženého rozdělení, získáme použitím  $i$ -té realizace každého marginálního rozdělení a přiřazením příslušné pravděpodobnosti.

Protože jsou všechny tyto veličiny generovány náhodně, ale jejich počet je konečný, mají téměř nulové korelace, nulové střední hodnoty a jednotkové rozptyly. Dále je zapotřebí pro korelační matici  $R$  určit její Choleského rozklad  $R = LL^T$ . Posledním krokem první fáze je výpočet parametrů později vstupujících do kubické transformace. To provedeme pomocí konkrétního systému rovnic pro případ prvních čtyř momentů.

Ve druhé, transformační, fázi aplikujeme na veličiny  $\mathbf{x}$  vygenerované v předchozí fázi transformaci  $\mathbf{y} = L\mathbf{x}$ , čímž získáme požadované korelace. A jako poslední transformaci provedeme dosazení získaných veličin do konkrétního systému rovnic jehož parametry jsme spočítali v předchozí fázi. Touto úpravou získáme požadované marginální momenty, ale mírně se pozmění korelace. Tím získáme nějakou aproximační chybu, kterou je možné zmenšovat iterováním lehce pozměněné druhé fáze.

Největší výhodou tohoto algoritmu je jeho rychlost. Na rozdíl od ostatních algoritmů se generuje pouze jedno marginální rozdělení v daném okamžiku a sdružené rozdělení je následně vytvořeno z těchto marginálních. Druhým rozdílem je způsob, kterým se tento algoritmus vyrovnává se změnami v zadaných statistických vlastnostech v průběhu transformací, které je mohou měnit. Jiný podobný algoritmus přepočítává korelační matici, aby byl dosažený výsledek co nejlepší, kdežto tento algoritmus přepočítává počáteční momenty. Tvůrci věří, že je tento přístup flexibilnější než v případě, kdy se upravují korelace.

Tento algoritmus je pro účely práce, kde se zabýváme jen předepsanými prvními a druhými momenty, příliš složitý.

## 2.2 Pravidelné rozložení scénářů

### 2.2.1 Shoda prvních momentů

Nejprve se omezme na nejjednodušší případ úlohy (1.2), kdy  $p^s = \frac{1}{S} \forall s = 1, \dots, S$  a kdy nás zajímají pouze první momenty. To znamená řešit úlohu

$$\sum_{s=1}^S \frac{1}{S} \omega_i^s = \mu_1(i), \quad i = 1, \dots, n. \quad (2.1)$$

V takovémto případě jsou  $\omega_i^s$  vzájemně nezávislé pro  $\forall i = 1, \dots, n$  a je možné řešit úlohu pro každé  $i = 1, \dots, n$  zvlášť.

Jednou z možností, jak úlohu (2.1) řešit, je rozložit  $\omega_i^1, \dots, \omega_i^S$  pravidelně a symetricky okolo  $\mu_1(i)$ . Navíc bez újmy na obecnosti předpokládejme, že  $\omega_i^1 < \omega_i^2 < \dots < \omega_i^S, S > 1$ . (Pokud  $S = 1$ , pak  $\omega_i^1 = \mu_1(i)$ .) To znamená, že chceme, aby platilo následující:

$$\begin{aligned} \omega_i^{j+1} - \omega_i^j &= \frac{\omega_i^S - \omega_i^1}{S-1}, \quad j = 1, \dots, S-1 \\ \omega_i^k - \mu_1(i) &= -(\omega_i^{S-k+1} - \mu_1(i)), \quad k = 1, \dots, S \\ \omega_i^j &\neq \omega_i^k, \quad \forall j, k = 1, \dots, S, \quad j \neq k. \end{aligned} \quad (2.2)$$

Označme si

$$a_i^s := \omega_i^s - \mu_1(i), \quad (2.3)$$

pak, aby byly splněny podmínky (2.2) platí, že

$$a_i^s = -a_i^{S-s+1} \quad (2.4)$$

$$a_i^s = a_i^1 \frac{S+1-2s}{S-1} \quad (2.5)$$

Pak platí, že

$$\begin{aligned} \mu_1(i) &= \sum_{s=1}^S \frac{1}{S} \omega_i^s = \frac{1}{S} \sum_{s=1}^S (a_i^s + \mu_1(i)) = \frac{1}{S} \sum_{s=1}^S \left( a_i^1 \frac{S+1-2s}{S-1} \right) + \frac{1}{S} \sum_{s=1}^S \mu_1(i) = \\ &= \frac{a_i^1 S(S+1)}{S(S-1)} - \frac{2a_i^1}{S(S-1)} \sum_{s=1}^S s + \frac{S\mu_1(i)}{S} = \\ &= \frac{a_i^1 S(S+1)}{S(S-1)} - \frac{2a_i^1 S(S+1)}{2S(S-1)} + \mu_1(i) = \mu_1(i) \end{aligned}$$

To znamená, že pro všechna  $i = 1, \dots, n$  a námi zvolená  $\omega_i^1 < \mu_1(i)$  nalezneme vhodné  $\omega_i^s, s = 2, \dots, S$  tak, aby splňovaly (2.1). Metodu je možné uplatnit i v případě, pokud jako inicializační scénář použijeme jiný, než ten nejmenší a budeme vědět kolikátý z nich to je. Výjimku tvoří pouze případ, kdy  $S$  je liché číslo a zvolený by byl scénář  $\omega_i^{\lceil \frac{S}{2} \rceil}$  jehož hodnota je rovna hodnotě  $\mu_1(i)$ .

Výsledné scénáře pak tvoří matice velikosti  $S * i$ , kterých je  $(S!)^i$ . Jednou takovou maticí je například matice  $M$  s prvky  $\omega_i^s, s = 1, \dots, S, i = 1, \dots, n$ , kde  $s$  odpovídá řádce a  $i$  odpovídá sloupci. Ostatní matice získáme permutováním prvků uvnitř sloupců matice  $M$ .

Pokud bychom považovali matici  $M$  za jediné řešení řešení, přidali bychom závislost mezi jednotlivé  $s$ -té složky.

## 2.2.2 Shoda prvních momentů a rozptylů

Nyní se zabýváme úlohou (1.2) tak, že nás budou zajímat první momenty, rozptyly a  $p^s = \frac{1}{S} \forall s = 1, \dots, S$ . Tím získáme úlohu

$$\begin{aligned} \sum_{s=1}^S \frac{1}{S} \omega_i^s &= \mu_1(i), \quad i = 1, \dots, n \\ \sum_{s=1}^S \frac{1}{S} (\omega_i^s - \mu_1(i))^2 &= \text{var}(i), \quad i = 1, \dots, n. \end{aligned} \quad (2.6)$$

Tuto úlohu můžeme opět řešit pomocí pravidelného a symetrického rozložení scénářů okolo  $\mu_1(i)$ , to znamená požadovat stejné podmínky jako (2.2) a definovat (2.3).

Shody prvních momentů dosáhneme pro každé  $\omega_i^1 < \mu_1(i)$  jako v předchozím odstavci.

Pro shodu rozptylů platí, že

$$\begin{aligned} \text{var}(i) &= \sum_{s=1}^S \frac{1}{S} (\omega_i^s - \mu_1(i))^2 = \frac{1}{S} \sum_{s=1}^S (a_i^s + \mu_1(i) - \mu_1(i))^2 = \\ &= \frac{1}{S} \sum_{s=1}^S (a_i^s)^2 = \frac{1}{S} \sum_{s=1}^S \left( a_i^1 \frac{S+1-2s}{S-1} \right)^2 = \frac{(a_i^1)^2}{S(S-1)^2} \sum_{s=1}^S (S+1-2s)^2 = \\ &= \frac{(a_i^1)^2}{S(S-1)^2} \sum_{s=1}^S ((S^2 + 2S + 1) + (4s^2 - 4sS - 4s)) = \\ &= \frac{(a_i^1)^2}{S(S-1)^2} \left( \sum_{s=1}^S (S^2 + 2S + 1) + \sum_{s=1}^S (4s^2) + \sum_{s=1}^S (-4s(S+1)) \right) = \\ &= \frac{(a_i^1)^2}{S(S-1)^2} \left( S(S^2 + 2S + 1) + 4 \frac{S(S+1)(2S+1)}{6} - 4 \frac{S(S+1)}{2} (S+1) \right) \\ &= \frac{1}{3} \frac{S+1}{S-1} (a_i^1)^2. \end{aligned}$$

To znamená, že

$$a_i^1 = \sqrt{\frac{3(S-1) \text{var}(i)}{S+1}}, \quad (2.7)$$



a dosazením do (2.5) a (2.3) určíme všechny scénáře  $\omega_i^1, \dots, \omega_i^S$  tak, aby splňovaly (2.6).

## 2.3 Soustava nelineárních rovnic

Další ze způsobů řešení úlohy (1.2) zmiňuje R. Kouwenber v práci [11], kde se  $p^s = \frac{1}{S} \forall s = 1, \dots, S$  a se zabývá shodou pouze středních hodnot a kovariancí. To znamená, že řeší soustavu nelineárních rovnic o  $n * S$  neznámých následujícího tvaru

$$\frac{1}{S} \sum_{s=1}^S \omega_i^s = \mu_1(i), \forall i = 1, \dots, n$$

$$\frac{1}{S} \sum_{s=1}^S (\omega_i^s - \mu_1(i)) (\omega_j^s - \mu_1(j)) = \text{cov}_{i,j}, \forall i = 1, \dots, n-1, j = i+1, \dots, n,$$

kteřou lze za použití označení (2.3) převést na tvar

$$\frac{1}{S} \sum_{s=1}^S a_i^s = 0, \forall i = 1, \dots, n$$

$$\frac{1}{S} \sum_{s=1}^S a_i^s a_j^s = \text{cov}_{i,j}, \forall i = 1, \dots, n-1, j = i+1, \dots, n.$$

Nyní si rozmysleme kolik scénářů  $S$  můžeme hledat pro  $n$ -složkový vektor, abychom získali jednoznačné neparametrické řešení. Necht'  $\omega^s = (\omega_1^s, \dots, \omega_n^s)$ , kde  $n$  je známé, pak získáme  $n$  rovnic  $\frac{1}{S} \sum_{s=1}^S a_i^s = 0$  a  $\frac{1}{2} \frac{n!}{(n-2)!}$  rovnic  $\frac{1}{S} \sum_{s=1}^S a_i^s a_j^s = \text{cov}_{i,j}$ . Tento počet rovnic je zapotřebí ještě vydělit hodnotou  $n$ , protože každý  $\omega^s$  má  $n$  složek. Tím zjistíme, že

$$S = \lfloor 1 + \frac{n-1}{2} \rfloor,$$

což zároveň znamená, že

$$n = \begin{cases} 2S & \text{pro } n \text{ sudé} \\ 2S - 1 & \text{pro } n \text{ liché.} \end{cases}$$

## 2.4 Řešení pomocí cílového programování

**Definice 21.** (Vícekritériální programování) [5, 3.1.1] Řešení  $\hat{\mathbf{x}} \in \mathcal{X}$  je eficientní řešení úlohy „minimalizujte“ funkci  $\mathbf{f} = (f_1, \dots, f_K)$ ,  $K \geq 2$ ,  $f_k : \mathbb{R}^n \rightarrow \mathbb{R}^1$  na uzavřené množině  $\mathcal{X}$  (zkráceně „min“  $\mathbf{f}(\mathbf{x})$  na  $\mathcal{X}$ ), pokud neexistuje žádné  $\mathbf{x} \in \mathcal{X}$ , pro které  $\mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\hat{\mathbf{x}})$  a  $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\hat{\mathbf{x}})$ .

Jednou z metod nalezení eficientního řešení je cílové programování. Jeho hlavní ideou je získat řešení z  $\mathcal{X}$ , pro které výstup změřený jako

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_K(\mathbf{x}))^\top$$

je co nejbliže k vektoru délky  $K$  složenému z nejlepších dosažitelných řešení pro jednotlivé funkce

$$f_k^* = \min_{\mathbf{x} \in \mathcal{X}} f_k(\mathbf{x}), k = 1, \dots, K.$$

Vzdálenosti jsou definovány v prostoru hodnot funkce  $\mathbf{f}$  a podmnožiny  $\mathbb{R}^K$ , kde můžeme použít kteroukoli z  $L_p$  vzdáleností,  $1 \leq p < \infty$ . Jedná se tedy o minimalizační problém

$$\min_{\mathbf{x} \in \mathcal{X}} \|\mathbf{T}(\mathbf{f}^* - \mathbf{f}(\mathbf{x}))\|_p \quad (2.8)$$

s diagonální maticí  $\mathbf{T} = \text{diag}\{t_1, \dots, t_K\}$ ,  $t_k > 0 \forall k$ . Sjednocením obecné úlohy cílového programování (2.8) s úlohou (1.2) zjistíme, že

$$\begin{aligned} \mathbf{f}^* &= (\mu_k(1), \dots, \mu_k(n), \text{cov}_{1,j_1}, \dots, \text{cov}_{n-1,j_{n-1}})^\top, \\ \mathbf{f}(\boldsymbol{\omega}) &= \left( \sum_{s=1}^S p^s (\omega_1^s)^k, \dots, \sum_{s=1}^S p^s (\omega_n^s)^k, \sum_{s=1}^S p^s (\omega_1^s - \mu_1(1)) (\omega_{j_1}^s - \mu_1(j_1)), \dots, \right. \\ &\quad \left. \sum_{s=1}^S p^s (\omega_{n-1}^s - \mu_1(n-1)) (\omega_{j_{n-1}}^s - \mu_1(j_{n-1})) \right)^\top, \\ \mathbf{T} &= \text{diag}\{\alpha_{1,k}, \dots, \alpha_{n,k}, \beta_{1,j_1}, \dots, \beta_{n-1,j_{n-1}}\} \\ &\text{pro } k = 1, \dots, m, j_i = i + 1, \dots, n. \end{aligned}$$

To tedy znamená, že řešení bude nalezeno jako řešení úlohy

$$\begin{aligned} \min_{\boldsymbol{\omega}} \|\text{diag}\{\alpha_{lk}, \beta_{ij}\} \left( (\mu_k(l), \text{cov}_{i,j})^\top - \right. \\ \left. - \left( \sum_{s=1}^S p^s (\omega_l^s)^k, \sum_{s=1}^S p^s (\omega_i^s - \mu_1(i)) (\omega_j^s - \mu_1(j)) \right)^\top \right)\|_p, \\ k = 1, \dots, m, l = 1, \dots, n, i = 1, \dots, n-1, j = i+1, \dots, n \end{aligned} \quad (2.9)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2, \dots, \omega_n^s \geq c_n, p^s \geq 0, s = 1, \dots, S, \sum_{s=1}^S p^s = 1$ , kde váhy  $\alpha_*, \beta_*$  odpovídají kvalitě a důležitosti dat a volba  $S$  závisí na uživateli.

# Kapitola 3

## Charakter úlohy (2.9)

Hlavní výhodou této formulace úlohy je, že optimální hodnota je 0, pokud jsou data konzistentní a  $S$  je dostatečně velké, avšak řešení je rovněž dobrou reprezentací dat, pokud jsou nekonzistentní.

Nekonzistentnost se může objevit například, pokud informace o hodnotách momentů získáme z různých zdrojů nebo pokud jsou implicitní specifika nekonzistentní s explicitními. Často je racionální určit hodnotu některých statistických vlastností na základě úsudku experta a jiné na základě empirické analýzy.

Jako příklad si představme, že si zadavatel přeje odhad kovariance a středních kvadratických odchylek určit z empirických dat, ale pro střední hodnotu se rozhodne využít svého úsudku o její hodnotě. Střední kvadratická odchylka a kovariance jsou tedy založeny na empirické střední hodnotě, která je nejspíše jiná než jakou odhaduje zadavatel. Z tohoto důvodu specifikované statistické vlastnosti daného rozdělení nemusí být vnitřně konzistentní. Dokonce se může stát, že takovéto rozdělení ani nemusí existovat.

Jako druhý příklad uvažme problém, který pokrývá dvě časová období. Specifikujme střední kvadratickou odchylku  $\omega_1$  a střední kvadratickou odchylku součtu  $\omega_1 + \omega_2$ . Specifikováním těchto dvou kvadratických odchylek jsme ale zároveň řekli něco o korelaci v průběhu času. Pokud bychom teď explicitně specifikovali korelaci v průběhu času, nejspíše bychom tak získali dvě odlišné informace.

A jako poslední případ si uveďme situaci, kdy je kovariance specifikována podle úsudku manažera a druhý obecný moment je vypočítán z dat. V takovém případě opět získáváme nekonzistentnost, jelikož kovariance je druhým centrálním momentem. V tomto případě je však možné relativně jednoduše zjistit, jestli se nekonzistentnost objevila nebo ne, i když je jen málo pravděpodobné, že tyto dva údaje jsou konzistentní. Z definic druhého obecného a centrálního momentu a za předpokladu  $\text{var } X < \infty$  lze jednoduše odvodit vztah  $\text{var } X = \mathbf{E} X^2 - (\mathbf{E} X)^2$  neboli rozptyl je druhý obecný moment snížený o druhou mocninu střední hodnoty.

Algoritmická metoda řešení naší optimalizační úlohy zvládne nekonzistentnost statistických vlastností, ale nedosáhne nulové výsledné hodnoty. V takovém případě se musí manažer rozhodnout, zda se spokojí s řešením, které se ohledně statistických vlastností shoduje s původním pouze přibližně, nebo přehodnotí zadané specifikace a změní je. Manažer může dát každé specifikaci takovou váhu  $\alpha_*, \beta_*$ , aby dosáhl korektního kompromisu mezi dříve diskutovanými možnostmi.

Náročnost numerického řešení tohoto problému roste jak se zvětšujícím se

počtem dat, tak se zvětšující se dimenzí vektoru. Náročnost rovněž roste s množstvím momentů, na které máme požadavek shody.

### 3.1 Speciální volba rozměrů úlohy

Nyní se budeme věnovat problému, kdy  $n = 2$ ,  $p^s = \frac{1}{S} \forall s$ ,  $p = 1, 2$  a  $k = 1, 2$ . To tedy znamená, že budeme mít dvourozměrná data, všechny scénáře budou stejně pravděpodobné, normu zvolíme  $L_1$  a  $L_2$  a budeme požadovat pouze shodu středních hodnot, druhých momentů a případně kovariance. V další části si ukážeme i úlohu, ve které budeme požadovat shodu rozptylů.

Omezením pravděpodobností na stejné hodnoty, tedy konstanty, se problém velmi zjednoduší. Pokud bychom to neudělali, měli bychom dvakrát více proměnných a tedy podstatně složitější a výpočetně daleko náročnější problém. Navíc by ani nebylo možné udělat některé další výhodné úpravy a přeformulování, která umožňují jednodušší výpočty.

Volbou těchto výchozích podmínek lze náš problém přeformulovat na hledání hodnot vektorových dvojic  $(\omega_1^s, \omega_2^s)$ ,  $s = 1, \dots, S$  tak, že

$$\begin{aligned} \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^k &= \mu_k(1) \text{ pro } k = 1, 2 \\ \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^k &= \mu_k(2) \text{ pro } k = 1, 2 \\ \sum_{s=1}^S \frac{1}{S} (\omega_1^s - \mu_1(1))(\omega_2^s - \mu_1(2)) &= \text{cov}_{1,2} \end{aligned}$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Úloha (2.9) se tak redukuje na úlohu, kterou lze v maticové formě napsat takto:

$$\min \left\| \begin{pmatrix} \alpha_{1,1} & 0 & 0 & 0 & 0 \\ 0 & \alpha_{2,1} & 0 & 0 & 0 \\ 0 & 0 & \alpha_{1,2} & 0 & 0 \\ 0 & 0 & 0 & \alpha_{2,2} & 0 \\ 0 & 0 & 0 & 0 & \beta_{1,2} \end{pmatrix} \begin{pmatrix} \mu_1(1) - \sum_{s=1}^S \frac{1}{S} (\omega_1^s) \\ \mu_1(2) - \sum_{s=1}^S \frac{1}{S} (\omega_2^s) \\ \mu_2(1) - \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 \\ \mu_2(2) - \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 \\ COV \end{pmatrix} \right\|_p \quad (3.1)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ , kde

$$COV = \text{cov}_{1,2} - \sum_{s=1}^S \frac{1}{S} (\omega_1^s - \sum_{s=1}^S \frac{1}{S} (\omega_1^s)) (\omega_2^s - \sum_{s=1}^S \frac{1}{S} (\omega_2^s)).$$

## 3.2 Volba metriky

### 3.2.1 $L_1$ norma

#### Shoda středních hodnot

V nejjednodušším případě budeme požadovat pouze shodu středních hodnot (tedy  $k = 1$ ). Tím se úloha (3.1) mění na úlohu následujícího tvaru:

$$\min \left\| \begin{pmatrix} \alpha_{1,1} & 0 \\ 0 & \alpha_{2,1} \end{pmatrix} \begin{pmatrix} \mu_1(1) - \sum_{s=1}^S \frac{1}{S} \omega_1^s \\ \mu_1(2) - \sum_{s=1}^S \frac{1}{S} \omega_2^s \end{pmatrix} \right\|_p \quad (3.2)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Metriku zvolíme  $L_1$ , tím se úloha (3.2) mění na

$$\begin{aligned} & \min \left| \alpha_{1,1} \left( \mu_1(1) - \sum_{s=1}^S \frac{1}{S} \omega_1^s \right) \right| + \left| \alpha_{2,1} \left( \mu_1(2) - \sum_{s=1}^S \frac{1}{S} \omega_2^s \right) \right| = \\ & = \min \alpha_{1,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \right| + \alpha_{2,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \right| \end{aligned} \quad (3.3)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Absolutní hodnota je na  $\mathbb{R}$  konvexní funkce,  $\frac{1}{S} \sum_{s=1}^S \omega_i^s - \mu_1(i)$  pro  $i = 1, 2$  je afinní zobrazení dané předpisem  $(\frac{1}{S}, \dots, \frac{1}{S}) \begin{pmatrix} \omega_i^1 \\ \vdots \\ \omega_i^S \end{pmatrix} - \mu_1(i)$  a podle Věty 4 je tato funkce konvexní. Také vážený součet konvexních funkcí je konvexní při libovolných nezáporných vahách, jak říká Věta 2, proto je minimalizační úloha (3.3) úlohou konvexní.

Úlohu (3.3) lze přetransformovat na úlohu lineárního programování.

**Definice 22.** [7, Kapitola 1.1] *Úloha lineárního programování je úloha minimalizovat  $f(\mathbf{x})$  na množině řešení soustavy rovnic a nerovnic*

$$\begin{aligned} g_k(\mathbf{x}) &\leq 0, \quad k = 1, 2, \dots, m \\ h_j(\mathbf{x}) &= 0, \quad j = 1, 2, \dots, p \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned}$$

*přičemž všechny funkce  $f, g_k, h_j$  jsou lineární.*

Označme si

$$\begin{aligned} \gamma_1 &:= \left| \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \right| \\ \gamma_2 &:= \left| \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \right|. \end{aligned}$$

Cílem je minimalizovat  $\alpha_{1,1}\gamma_1 + \alpha_{2,1}\gamma_2$

$$\begin{aligned} \text{za podmíněk } & -\gamma_1 \leq \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \leq \gamma_1 \\ & -\gamma_2 \leq \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \leq \gamma_2 \\ & \omega_1^s \geq c_1, \omega_2^s \geq c_2. \end{aligned}$$

To znamená řešit úlohu lineárního programování ve tvaru:

$$\begin{aligned} & \min \alpha_{1,1}\gamma_1 + \alpha_{2,1}\gamma_2 \\ \text{za podmíněk } & \gamma_1 + \frac{1}{S} \sum_{s=1}^S \omega_1^s \geq \mu_1(1) \\ & \gamma_1 - \frac{1}{S} \sum_{s=1}^S \omega_1^s \geq -\mu_1(1) \\ & \gamma_2 + \frac{1}{S} \sum_{s=1}^S \omega_2^s \geq \mu_1(2) \\ & \gamma_2 - \frac{1}{S} \sum_{s=1}^S \omega_2^s \geq -\mu_1(2) \\ & \omega_1^s \geq c_1 \\ & \omega_2^s \geq c_2. \end{aligned} \tag{3.4}$$

Jak bylo zmíněno na začátku kapitoly (3.1), úlohu (3.3) by nebylo možné přeformulovat na úlohu lineárního programování (3.4), pokud by pravděpodobnosti byly optimalizované proměnné.

### Shoda vyšších momentů

Ve druhém případě ponecháme metriku  $L_1$ , ale tentokrát budeme požadovat shodu rozptylů. Tím získáme optimalizační úlohu následujícího tvaru:

$$\begin{aligned} & \min \left| \alpha_{1,2} \left( \text{var}(1) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \right) \right| + \\ & + \left| \alpha_{2,2} \left( \text{var}(2) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \right) \right| \end{aligned} \tag{3.5}$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Tato úloha není konvexní. Uvažujme, že hledáme pouze dva scénáře ( $S = 2$ ). Po jednoduchých úpravách získáme úlohu  $\min \alpha_{1,2} \left| \text{var}(1) - \frac{1}{4}(\omega_1^1 - \omega_1^2)^2 \right| + \alpha_{2,2} \left| \text{var}(2) - \frac{1}{4}(\omega_2^1 - \omega_2^2)^2 \right|$  při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ , kde  $\text{var}(i) > 0$  je konstanta. Při zafixovaném  $\omega_i^1$  funkce  $\left| \text{var}(i) - \frac{1}{4}(\omega_i^1 - \omega_i^2)^2 \right|$  není konvexní, proto ani úloha (3.5) není konvexní.

Rovněž můžeme formulovat úlohu na shodu druhých momentů. Ta bude mít následující tvar:

$$\min \alpha_{1,2} \left| \mu_2(1) - \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 \right| + \alpha_{2,2} \left| \mu_2 - \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 \right| \quad (3.6)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Pokud bychom z pozorovaných dat znali pouze střední hodnotu a rozptyl byla by úloha na shodu druhých momentů přeformulována za pomoci vzorce  $\text{var } X = \text{E } X^2 - (\text{E } X)^2$  do tvaru

$$\begin{aligned} \min \alpha_{1,2} \left| \left( \text{var}(1) + (\mu_1(1))^2 \right) - \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 \right| + \\ + \alpha_{2,2} \left| \left( \text{var}(2) + (\mu_1(2))^2 \right) - \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 \right| \end{aligned} \quad (3.7)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Tyto úlohy rovněž nejsou konvexní. Jako jednoduchý příklad si představme situaci, kdy  $S = 1$ . V takovém případě získáme úlohu  $\min \alpha_{1,2} |a_1 - (\omega_1^1)^2| + \alpha_{2,2} |a_2 - (\omega_2^1)^2|$ , kde  $a_i = \mu_2(i)$  pro (3.6) a  $a_i = \text{var}(i) + (\mu_1(i))^2$ . Navíc v obou úlohách platí, že  $a_i > 0$  je konstanta. Funkce  $|a_i - (\omega_i^1)^2|$  není konvexní, proto ani úlohy (3.6) a (3.7) nejsou konvexní.

Řešením úlohy (3.5) a (3.6) potažmo (3.7) získáme data, která budou mít požadovaný druhý centrální nebo obecný moment. Může se však stát, že vygenerované scénáře nebudou původní data dobře reprezentovat, protože nepožadujeme shodu středních hodnot, díky které by byly ukotveny v prostoru na správném místě.

Proto, pokud chceme získat scénáře reprezentující data, je zapotřebí řešit následující funkci:

$$\begin{aligned} \min \alpha_{1,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \right| + \alpha_{2,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \right| + \\ + \alpha_{1,2} \left| \text{var}(1) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \right| + \\ + \alpha_{2,2} \left| \text{var}(2) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \right| \end{aligned} \quad (3.8)$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ . Takto formulovaná funkce zohledňuje první moment a rozptyl.

Úloha (3.8) není konvexní ze stejného důvodu jako (3.5), kterou obsahuje.

Navíc je možné (3.8) přeformulovat obdobně jako úlohu (3.4). Tím získáme

následující tvar:

$$\min \alpha_{1,1}\gamma_{1,1} + \alpha_{2,1}\gamma_{2,1} + \alpha_{1,2}\gamma_{1,2} + \alpha_{2,2}\gamma_{2,2}$$

$$\text{za podmínek } \gamma_{1,1} + \frac{1}{S} \sum_{s=1}^S \omega_1^s \geq \mu_1(1)$$

$$\gamma_{1,1} - \frac{1}{S} \sum_{s=1}^S \omega_1^s \geq -\mu_1(1)$$

$$\gamma_{2,1} + \frac{1}{S} \sum_{s=1}^S \omega_2^s \geq \mu_1(2)$$

$$\gamma_{2,1} - \frac{1}{S} \sum_{s=1}^S \omega_2^s \geq -\mu_1(2)$$

$$\gamma_{1,2} + \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \geq \text{var}(1)$$

$$\gamma_{1,2} - \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \geq -\text{var}(1)$$

$$\gamma_{2,2} + \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \geq \text{var}(2)$$

$$\gamma_{2,2} - \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \geq -\text{var}(2)$$

$$\omega_1^s \geq c_1$$

$$\omega_2^s \geq c_2.$$

(3.9)

Úloha (4.1) už ale není úlohou lineárního programování, protože některé z funkcí v podmínkách nejsou lineární, a je výpočetně i časově náročnější.

### 3.2.2 $L_2$ norma

Nyní zvolme  $p = 2$ . V tomto případě si ukážeme, jakou úlohu získáme, pokud budeme požadovat shodu středních hodnot, rozptylů i kovariancí, což znamená



aplikovat normu  $p = 2$  na úlohu (3.1). Úloha (3.1) má tedy následující tvar:

$$\begin{aligned}
& \min \left\{ \left| \alpha_{1,1} \left( \mu_1(1) - \frac{1}{S} \sum_{s=1}^S \omega_1^s \right) \right|^2 + \right. \\
& \quad + \left| \alpha_{2,1} \left( \mu_1(2) - \frac{1}{S} \sum_{s=1}^S \omega_2^s \right) \right|^2 + \\
& \quad + \left| \alpha_{1,2} \left( \text{var}(1) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \right) \right|^2 + \\
& \quad + \left| \alpha_{2,2} \left( \text{var}(2) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \right) \right|^2 + \\
& \quad \left. + \left| \beta_{1,2} \left( \text{cov}_{1,2} - \sum_{s=1}^S \frac{1}{S} (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s) (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s) \right) \right|^2 \right\}^{\frac{1}{2}} \quad (3.10)
\end{aligned}$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Nás ale nezajímá přesná výsledná hodnota minimalizační úlohy, ale pouze body minima, proto není nutné a z důvodu větší výpočetní náročnosti problému ani vhodné v úloze ponechávat druhé odmocniny, které tam mají podle definice  $L_p$ -normy být. Rovněž není nutné vzhledem ke druhé mocnině ponechávat ve formulaci absolutní hodnoty. Tedy můžeme řešení úlohy (3.9) hledat ekvivalentně pomocí metody nejmenších čtverců jako řešení úlohy:

$$\begin{aligned}
& \min \alpha_{1,1}^2 \left( \mu_1(1) - \frac{1}{S} \sum_{s=1}^S \omega_1^s \right)^2 + \\
& \quad + \alpha_{2,1}^2 \left( \mu_1(2) - \frac{1}{S} \sum_{s=1}^S \omega_2^s \right)^2 + \\
& \quad + \alpha_{1,2}^2 \left( \text{var}(1) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \right)^2 + \\
& \quad + \alpha_{2,2}^2 \left( \text{var}(2) - \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \right)^2 + \\
& \quad + \beta_{1,2}^2 \left( \text{cov}_{1,2} - \sum_{s=1}^S \frac{1}{S} (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s) (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s) \right)^2 \quad (3.11)
\end{aligned}$$

při podmínkách  $\omega_1^s \geq c_1, \omega_2^s \geq c_2$ .

Tato úloha opět není konvexní. Znovu si jako příklad uveďme situaci, kdy  $S = 2$ . Tímto omezením získáme úlohu, která bude obsahovat funkci  $(\text{var}(i) - \frac{1}{4}(\omega_i^1 + \omega_i^2)^2)^2$ , kde  $\text{var}(i)$  je nezáporná konstanta. Tato funkce není konvexní, proto ani úloha (3.9) není konvexní.

Použitím této transformace náš problém přechází na úlohu kvadratického programování.

**Definice 23.** [7, Kapitola 4.1] *Úloha kvadratického programování je úloha mini-*

malizovat  $f(\mathbf{x})$  na množině řešení soustavy nerovnic a rovnic

$$g_k(\mathbf{x}) \leq 0, \quad k = 1, 2, \dots, m,$$

$$h_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, p$$

$$\mathbf{x} \in \mathbb{R}^n,$$

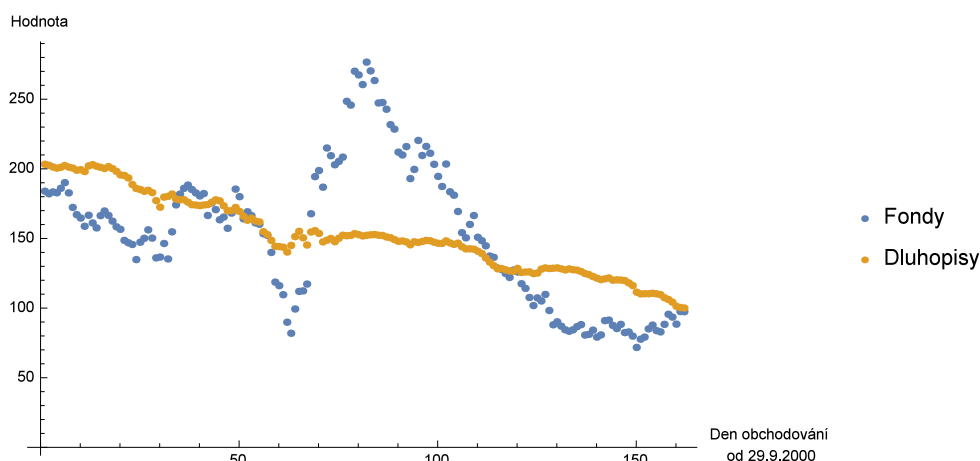
pokud  $f$  je kvadratická a  $g_k, h_j$  jsou lineární.

# Kapitola 4

## Praktická část

### 4.1 Data

Pro potřeby praktické části je využito měsíčních dat Conseq Invest akciový dluhopis A a Conseq Invest akciový fond A z období od 29. září 2000 do 1. dubna 2014, kdy byla data stažena, pocházejících z webové stránky [3]. Jednotlivé požadované statistické vlastnosti dat vstupující do výpočtů jako konstanty jsou určeny pomocí statistického softwaru R.



Obrázek 4.1: Použitá data

	Fond	Dluhopis
Minimum	71,74	100,06
Střední hodnota	152,53	152,30
Rozptyl	2573,0369	781,0882
Maximum	276,64	203,28

Tabulka 4.1: Popisné statistiky sledovaných veličin

Ještě poznamenejme, že všechny příklady jsou počítány pro  $\alpha_{n,k} = 1$ ,  $\beta_{n,k} = 1$  a ve všech grafech (s výjimkou kapitol 4.2.2, 4.3.2 a 4.4.3, kde jsou data umístěna na svých pozicích v reálných datech) jsou scénáře  $\omega_i^s$  přesunuty z původní pozice  $s$  na pozici  $s^* = \frac{162}{s+1}s$ , kde 162 je počet původních pozorování pro jednu složku.

## 4.2 Shoda prvních momentů

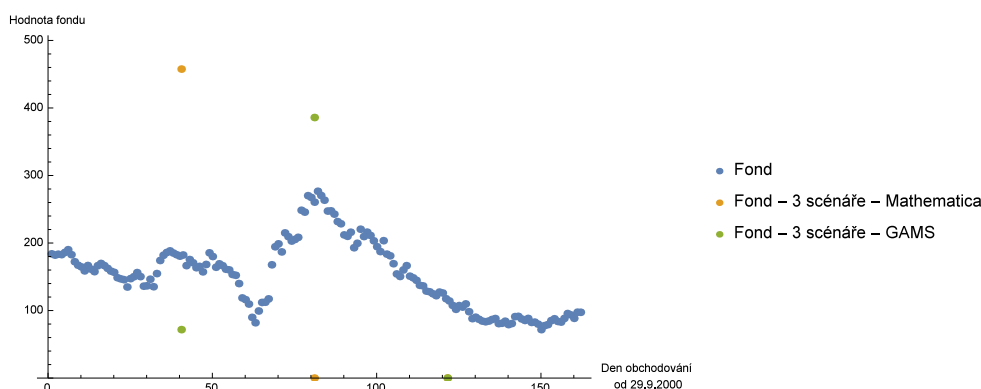
V této části si uvedeme dva možné přístupy k řešení úlohy (3.3). Prvním z nich je již dříve zmíněné převedení úlohy na úlohu lineárního programování (3.4) a generování nových scénářů tak, jako se řeší v celé práci, a druhým z nich je výběr z reálných dat.

### 4.2.1 Lineární programování

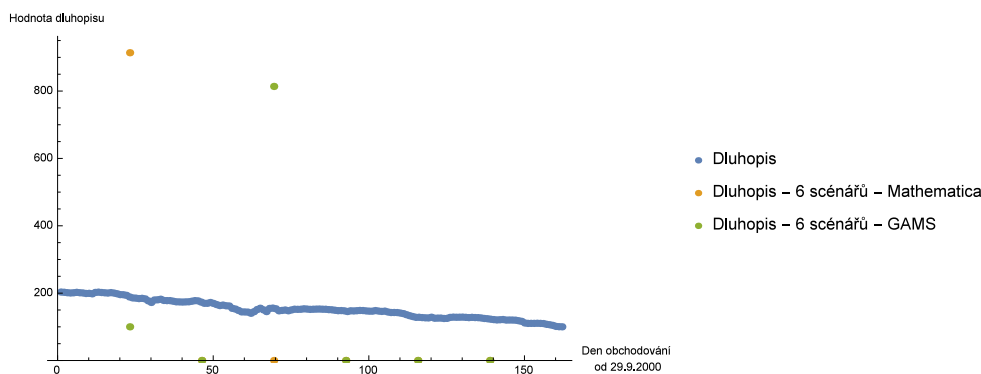
K řešení jsou použity dva výpočetní softwary. Prvním z nich je Wolfram Mathematica (dále jen Mathematica), kde je k výpočtu použita zabudovaná funkce lineárního programování. Druhým z nich je GAMS, ve které je potřeba úlohu lineárního programování programovat a je použit solver CPLEX.

#### Okrajové podmínky pro první scénář

Nejprve okrajovými podmínkami omezíme pouze první scénář, tzn.  $c_1 \leq \omega_1^1, c_2 \leq \omega_2^1, \omega_i^s \geq 0, s = 2, \dots, S$ .



Obrázek 4.2: Nevhodně spočítané 3 scénáře fondů



Obrázek 4.3: Nevhodně spočítaných 6 scénářů dluhopisů

Jak můžeme vidět na obrázcích (4.2) a (4.3), pokud počítáme úlohu (3.4) s omezujícími podmínkami pouze pro první scénář a podmínkou nezápornosti pro zbývající scénáře, tak dostaneme správné, ale nevhodné scénáře.

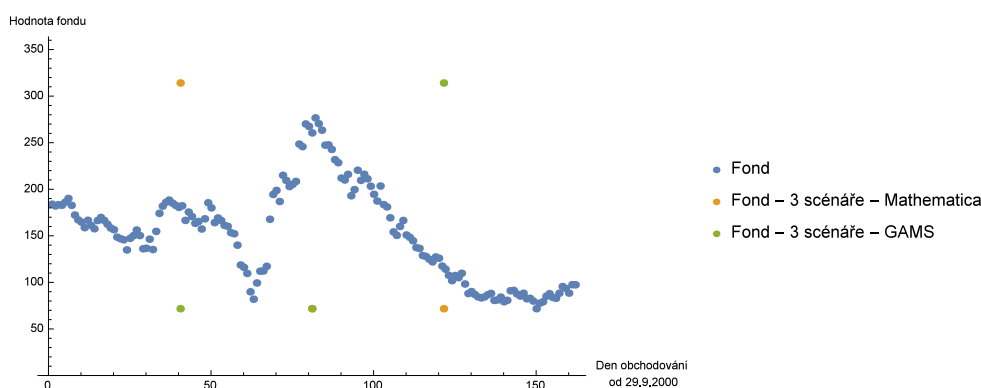
Software Mathematica hodnotu scénářů určí tak, že všechny scénáře kromě jednoho určí nulové a poslednímu scénáři určí vhodnou hodnotu. To je v pořádku pro  $S = 1, 2$ , ale pro  $S > 2$  tak získáme několik shodných nulových scénářů, což nechceme.

Software GAMS prvnímu scénáři určí minimální možnou hodnotu a ostatní kromě jednoho vypočítá jako nulové a hodnotu posledního scénáře určí vhodně. Slovem vhodně rozumějme tak, aby scénáře dosáhly požadované střední hodnoty. Takovéto řešení je vhodné pouze pro  $S = 1, 2, 3$ , ale pro  $S > 3$  opět získáme několik nechtěných shodných nulových scénářů.

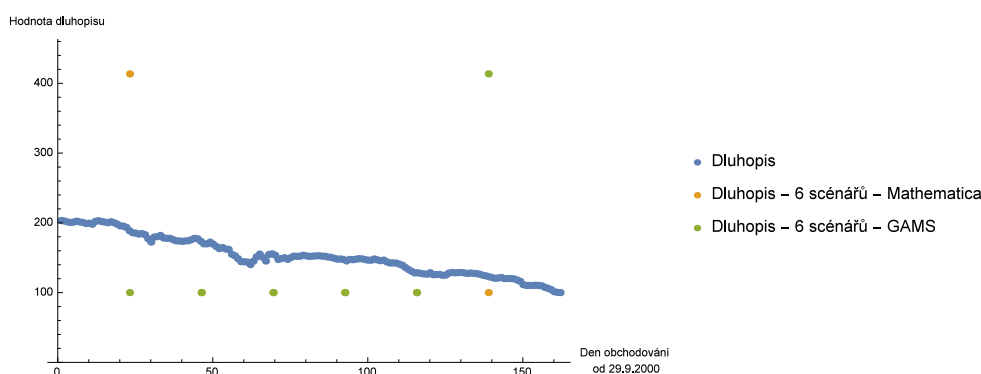
Proto je nutné úlohu doplnit o extrémální podmínky pro všechny scénáře, čímž zabráníme nulovým scénářům.

### Okrajové podmínky pro všechny scénáře

Přidáním okrajových podmínek pro všechny scénáře, tzn.  $c_1 \leq \omega_1^s, c_2 \leq \omega_2^s$ , k předchozí úloze získáme následující scénáře:



Obrázek 4.4: Proložení fondů 3 scénáři

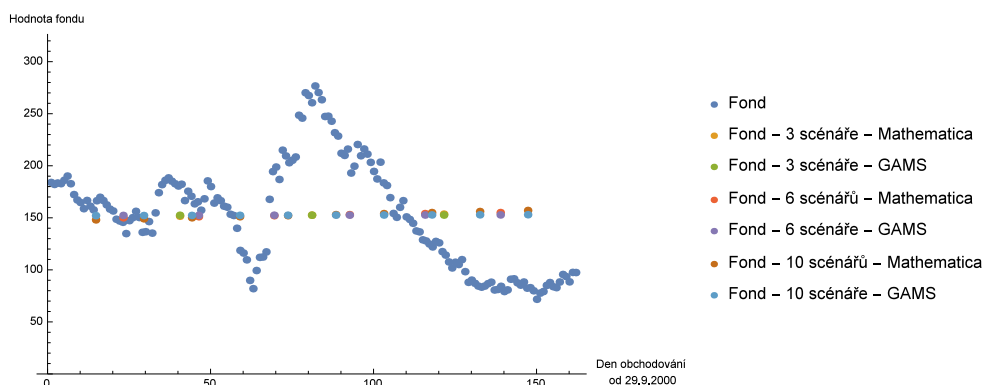


Obrázek 4.5: Proložení dluhopisů 6 scénáři

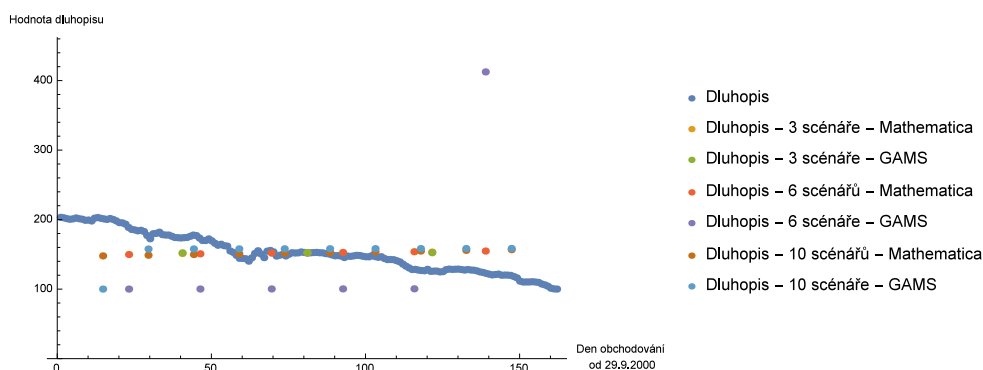
V obrázcích (4.4) a (4.5) můžeme nahlédnout, že v tomto případě, již máme  $S$  scénářů, které jsou nenulové, ale i přesto jsme získali v případech, kdy  $S > 3$ , opět shodné scénáře. Proto je zapotřebí přidat další podmínky zajišťující, aby žádné dva scénáře nemohly být stejné.

## Okrajové a nerovnostní podmínky pro všechny scénáře

Přidáním dalších podmínek  $\omega_1^s \neq \omega_1^k, \omega_2^s \neq \omega_2^k, \forall s, k = 1, \dots, S, s \neq k$  k již platným podmínkám  $c_1 \leq \omega_1^s, c_2 \leq \omega_2^s$  získáme následující scénáře:



Obrázek 4.6: Proložení fondů scénářů



Obrázek 4.7: Proložení dluhopisů scénářů

Na obrázcích (4.6) a (4.7) lze vidět, již vhodně spočítané scénáře z hlediska jejich různosti.

Rovněž můžeme vidět, že software Mathematica jednotlivé scénáře rozložil ve všech případech blízko střední hodnoty tak, aby nebyly shodné, ale jejich rozdíl je téměř nepatrný.

Podle obrázku (4.6) i software GAMS určil jednotlivé scénáře obdobným způsobem, přičemž velikost odchylky jednotlivých scénářů je nutno nastavovat v kódu. V obrázku (4.7) ale můžeme pozorovat jeho nekontrolovatelnost a nepředvídatelnost. 3 scénáře pro dluhopisy ještě určil obdobným způsobem. Avšak při dalších výpočtech se zachoval jinak. Při výpočtu 6-ti scénářů 5 z nich umístil k minimální možné hodnotě a 6-tý vhodně dopočítal, kdežto při výpočtu 10-ti scénářů určil pouze 1 scénář k minimální povolené hodnotě a zbylých 9 vypočítal vhodně poblíž střední hodnoty.

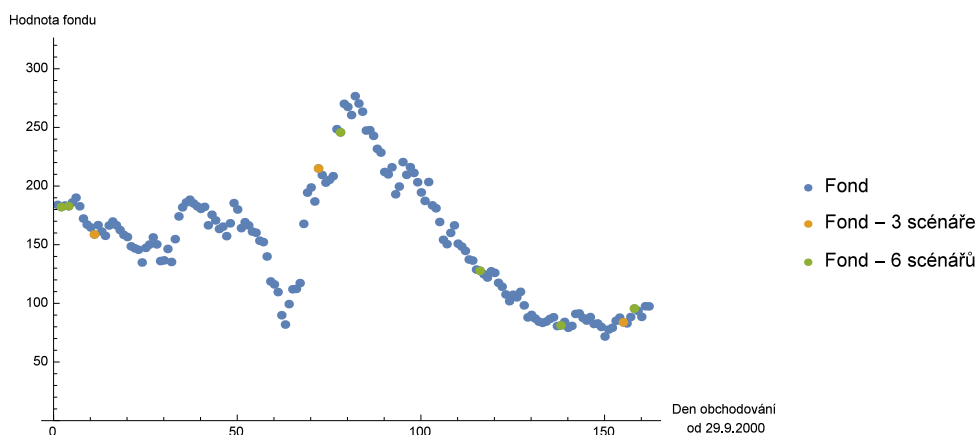
### 4.2.2 Výběr z reálných dat

K řešení pomocí vyhledávání v datech je zapotřebí mít data, pro která dané scénáře hledáme. Jakmile data máme je zapotřebí pro všechny možné  $S$ -tice

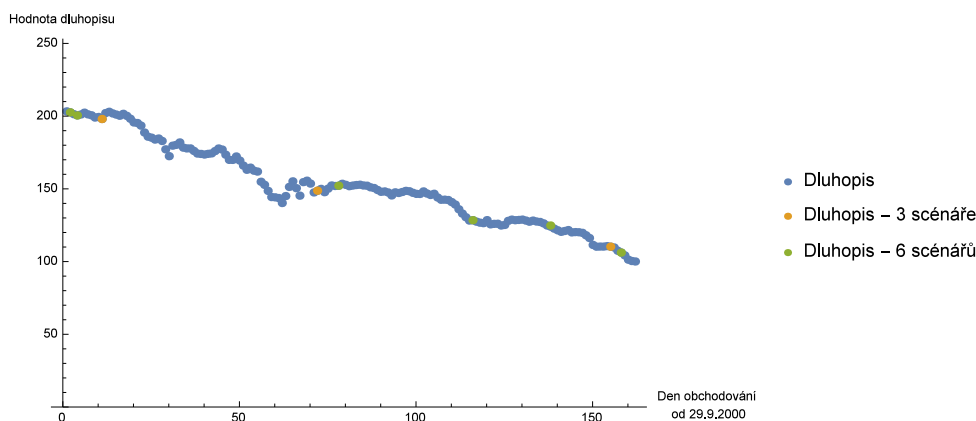
vyčíselit funkci

$$\alpha_{1,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \right| + \alpha_{2,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \right|$$

a nalézt minimum. Získáme tak tyto scénáře (jejich pozice odpovídá pozici v původních datech):



Obrázek 4.8: Nalezené scénáře fondu



Obrázek 4.9: Nalezené scénáře dluhopisu

### 4.2.3 Scénáře nalezené pro shodu prvních momentů

V následující tabulce (4.2) můžeme vidět, jak spolehlivé jednotlivé metody jsou.

Generováním nových scénářů jsme dosáhli absolutní shody s požadavky, což znamená, že hodnota  $\left| \frac{1}{S} \sum_{s=1}^S \omega_i^s - \mu_1(i) \right| = 0$  pro všechna  $i = 1, 2$ , proto hodnoty jednotlivých vah  $\alpha_*$  nemají na řešení vliv.

Jediný způsob řešení, kde váhy mohou řešení ovlivnit, je výběr z reálných dat. Avšak kvalita takového řešení je vždy přímo závislá na tvaru dat.

Rovněž lze nahlédnout, jak se projeví různé přístupy k rozmisťování scénářů na hodnotě rozptylu. Pomocí softwaru Mathematica jsme získali scénáře, které jsou vždy blízko střední hodnoty a rozptyly jsou malé.

Tohoto jsme dosáhli i pomocí softwaru GAMS v případech, kdy generoval obdobně. Avšak v případech generování 6-ti a 10-ti scénářů pro dlouhopisové složky už jsou rozptyly podstatně větší, přičemž pro 10 scénářů je hodnota rozptylu relativně nízká, protože pouze jedna hodnota není vhodně okolo střední hodnoty. Naopak při generování 6-ti scénářů je dosaženo relativně velkého rozptylu, protože žádná z vygenerovaných hodnot není v blízkosti střední hodnoty.

I.	II.	III.	IV.	V.	VI.	VII.
3	M	F	152,53	0	1	-2572,04
		D	152,3	0	1	-780,088
		F+D		0		-3352,128
	G	F	152,53	0	0,1	-2573,03
		D	152,3	0	0,1	-781,078
		F+D		0		-3354,108
	RD	F	152,52	-0,02	4328	1754,97
		D	152,34	0,04	1936,93	1155,84
		F+D		0,02		2910,81
6	M	F	152,53	0	3,5	-2569,54
		D	152,3	0	3,5	-777,588
		F+D		0		-3347,128
	G	F	152,53	0	0,035	-2573
		D	152,3	0	16249	-15467,9
		F+D		0		-18040,9
	RD	F	152,53	-0,01	3887,1	1314,06
		D	152,31	-0,99	1669,7	888,611
		F+D		-1		2202,671
10	M	F	152,53	0	9,167	-2563,87
		D	152,3	0	9,167	-771,922
		F+D		0		-3335,792
	G	F	152,53	0	0,092	-2572,95
		D	152,3	0	336,977	-444,111
		F+D		0		-6352,853

Tabulka 4.2: Popisné statistiky nalezených scénářů pro shodu prvních momentů



Znak	Vysvětlivka
I.	$S$
II.	Způsob řešení úlohy
III.	Složka
IV.	$\frac{1}{S} \sum_{s=1}^S \omega^s$
V.	$\frac{1}{S} \sum_{s=1}^S \omega^s - \mu_1$
VI.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2 - \left(\frac{1}{S} \sum_{s=1}^S \omega^s\right)^2$
VII.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2 - \left(\frac{1}{S} \sum_{s=1}^S \omega^s\right)^2 - \text{var}$
M	Mathematica
G	GAMS
RD	Výběr z reálných dat
F	Fond
D	Dluhopis

Tabulka 4.3: Vysvětlivky

## 4.3 Shoda druhých momentů

V této části si ukážeme možnosti řešení úlohy (3.6). Tentokrát máme opět dvě možnosti. První je nelineární programování v softwarech Mathematica a GAMS a druhou je výběr z reálných dat.

### 4.3.1 Nelineární programování

V případě nelineárního programování v softwaru Mathematica je řešena úloha (3.7) pomocí zabudované funkce Minimize. Zároveň je úloha řešena pro každé  $i = 1, 2$  zvlášť, což je možné díky nezávislosti řešení. Tento přístup se ukázal jako časově výrazně výhodnější.

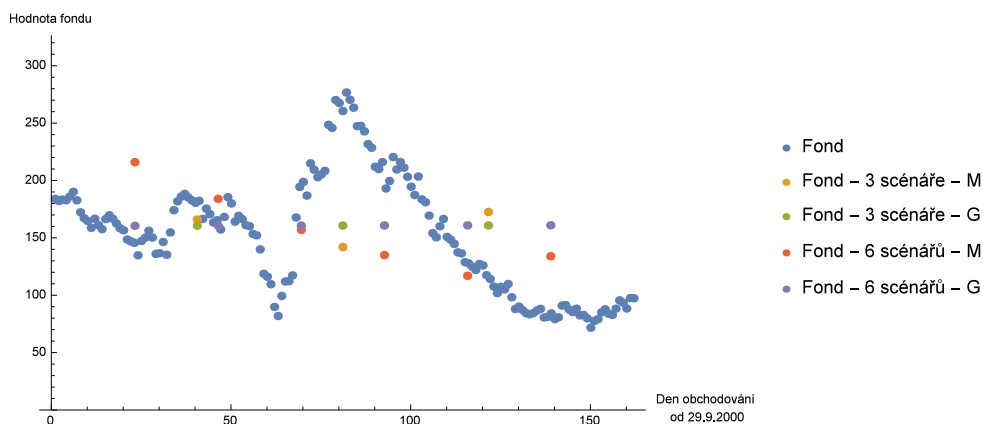
V softwaru GAMS je použito solveru CONOPT a je řešena následující úloha:

$$\begin{aligned}
& \min \alpha_{1,2} \gamma_{1,2} + \alpha_{2,2} \gamma_{2,2} \\
& \text{za podmínek } \gamma_{1,2} + \frac{1}{S} \sum_{s=1}^S (\omega_1^s)^2 \geq \text{var}(1) + (\mu_1(1))^2 \\
& \gamma_{1,2} - \frac{1}{S} \sum_{s=1}^S (\omega_1^s)^2 \geq -(\text{var}(1) + (\mu_1(1))^2) \\
& \gamma_{2,2} + \frac{1}{S} \sum_{s=1}^S (\omega_2^s)^2 \geq \text{var}(2) + (\mu_1(2))^2 \\
& \gamma_{2,2} - \frac{1}{S} \sum_{s=1}^S (\omega_2^s)^2 \geq -(\text{var}(2) + (\mu_1(1))^2) \\
& \omega_1^s \geq c_1 \\
& \omega_2^s \geq c_2 \\
& \omega_1^s \neq \omega_1^k, \forall s, k = 1, \dots, S, s \neq k \\
& \omega_2^s \neq \omega_2^k, \forall s, k = 1, \dots, S, s \neq k.
\end{aligned}$$

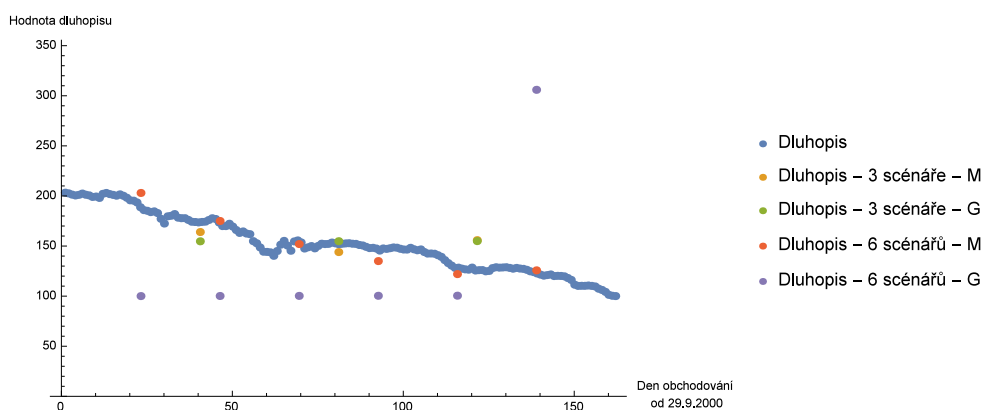
Opět je potřeba si dát pozor na dostatečné znění všech okrajových podmínek,

tak jako bylo demonstrováno v kapitole (4.2.1), abychom nezískali několik nevhodných shodných scénářů.

Získáme tak následující scénáře:



Obrázek 4.10: Nalezené scénáře fondu



Obrázek 4.11: Nalezené scénáře dluhopisu

### 4.3.2 Výběr z reálných dat

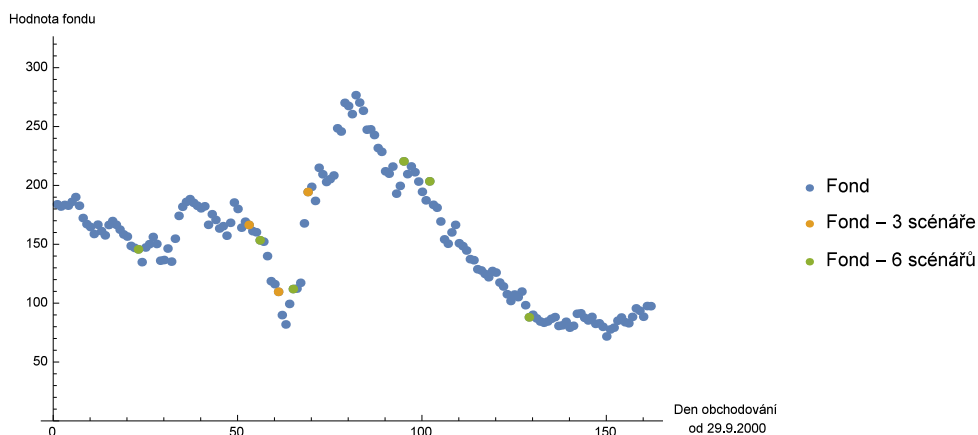
Opět je zapotřebí mít reálná dat, pro všechny  $S$ -tice vyčíslit funkci

$$\alpha_{1,2} \left| \frac{1}{S} \sum_{s=1}^S (\omega_1^s)^2 - \left( (\mu_1(1))^2 + \text{var}(1) \right) \right| +$$

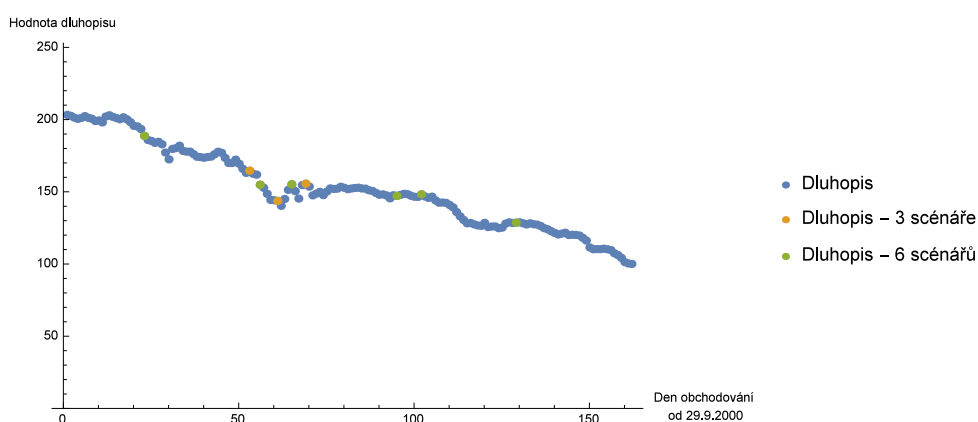
$$+ \alpha_{2,2} \left| \frac{1}{S} \sum_{s=1}^S (\omega_2^s)^2 - \left( (\mu_1(2))^2 + \text{var}(2) \right) \right|$$

a nalézt minimum.

Tímto postupem pro shodu druhých momentů nalezneme následující scénáře:



Obrázek 4.12: Nalezené scénáře fondu



Obrázek 4.13: Nalezené scénáře dluhopisu

### 4.3.3 Scénáře nalezené pro shodu druhých momentů

Jak můžeme v tabulce (4.4) vidět, nalezené scénáře již nemají správné střední hodnoty, což jsme ani nepožadovali. Požadovaných druhých momentů a tedy přesné shody bylo dosaženo při generování nových scénářů, což opět znamená, že hodnoty vah  $\alpha_*$  nemají na řešení vliv. Opět pouze u výběru z reálných dat mohou váhy ovlivnit kvalitu řešení, ale v přímé závislosti na tvaru dat.

I.	II.	III.	IV.	V.	VI.	VII.	VIII.
3	M	F	160,2	7,67	25838,4	0	259,484
		D	154,6	2,3	23976,4	0	101,172
		F+D		9,97		0	
	G	F	160,7	8,17	25838,3	0	0,01
		D	154,8	2,5	23976,4	0	0,01
		F+D		10,67		0	
	RD	F	156,8	4,27	25836,9	-1,5	1866,6
		D	154,6	2,3	23976,7	0,3	109,714
		F+D		6,57		-1,2	
6	M	F	157,166	-41,03	25838,4	0	1364,58
		D	152,124	-43,1	23976,4	0	1001,67
		F+D		-84,13		0	
	G	F	160,7	8,17	25838,3	0	0,035
		D	134,5	-17,8	23976,4	0	7050,59
		F+D		-9,63		0	
	RD	F	153,8	1,27	25838,5	0,1	2609,13
		D	153,8	1,5	23976,4	0	386,331
		F+D		2,77		0,1	

Tabulka 4.4: Popisné statistiky nalezených scénářů pro shodu druhých momentů

Znak	Vysvětlivka
I.	$S$
II.	Způsob řešení úlohy
III.	Složka
IV.	$\frac{1}{S} \sum_{s=1}^S \omega^s$
V.	$\frac{1}{S} \sum_{s=1}^S \omega^s - \mu_1$
VI.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2$
VII.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2 - ((\mu_1)^2 + \text{var})$
VIII.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2 - (\frac{1}{S} \sum_{s=1}^S \omega^s)^2$
M	Mathematica
G	GAMS
RD	Výběr z reálných dat
F	Fond
D	Dluhopis

Tabulka 4.5: Vysvětlivky

## 4.4 Shoda středních hodnot a rozptylů

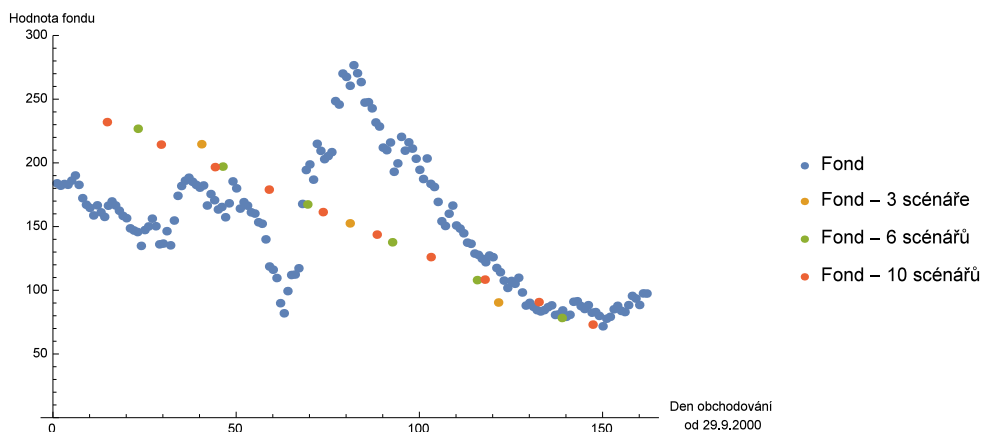
Při řešení úlohy s podmínkou shody středních hodnot a rozptylů máme opět několik možností. Můžeme generovat scénáře v softwarech GAMS a Mathematica, hledat nové scénáře metodou pravidelného a symetrického rozložení okolo střední hodnoty nebo opět vybírat z reálných dat.

### 4.4.1 Pravidelné rozložení scénářů

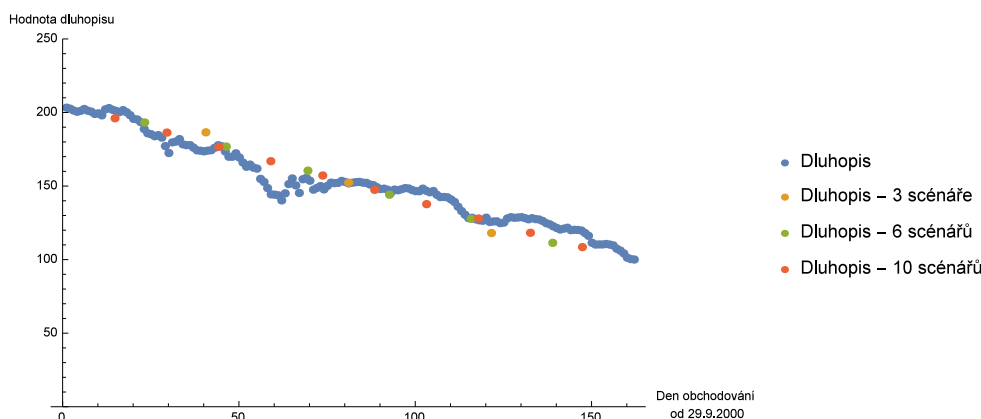
Nyní si numericky demonstrováme způsob řešení úlohy (1.2) pro shodu prvních momentů a rozptylů pomocí metody pravidelného rozložení scénářů okolo střední hodnoty pro hledání 3, 6 a 10 scénářů.

Postup výpočtu je následující:

- zvolit počet hledaných scénářů, tzn.  $S$
- určit hodnoty  $a_i^1$  z rovnice (2.7)
- vypočítat hodnoty  $a_i^s$  pomocí vzorce (2.5).



Obrázek 4.14: Proložení fondů scénáři



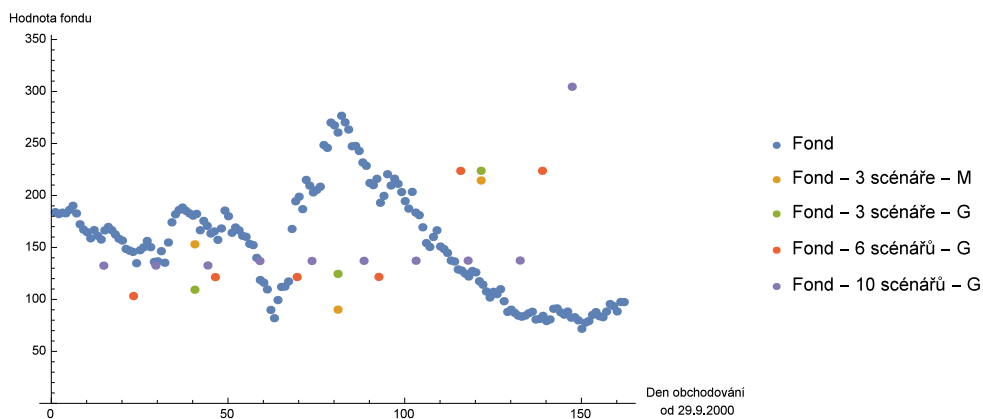
Obrázek 4.15: Proložení dluhopisů scénáři

Na obrázcích (4.2) a (4.3) můžeme vidět, že tento přístup zajišťuje pravidelné rozložení scénářů, které jsou symetricky rozloženy okolo střední hodnoty, a i minimální a maximální hodnoty dat nejsou překročeny.

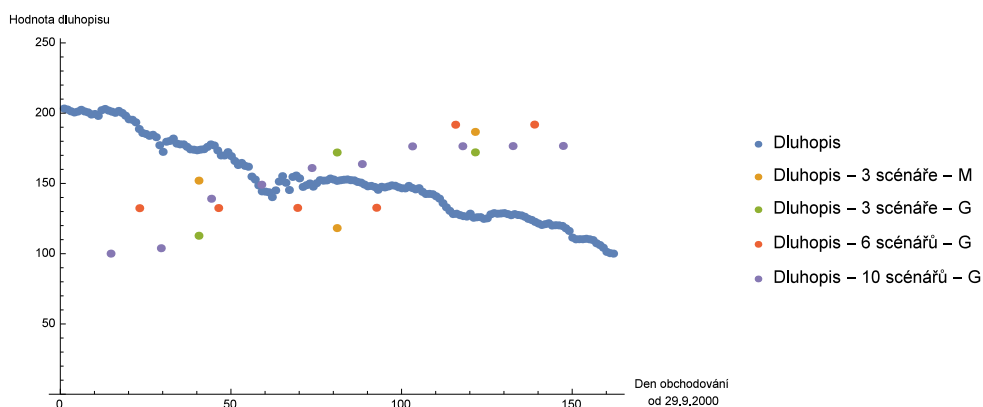
## 4.4.2 Nelineární programování

V softwaru Mathematica budeme scénáře generovat pomocí funkce `Minimize` jako řešení úlohy (3.8) a opět pro každé  $i = 1, 2$  zvlášť a pomocí softwaru GAMS budeme řešit úlohu nelineárního programování (3.9) za použití solveru CONOPT.

I v těchto příkladech generování scénářů musíme ošetřit všechny minimální a nerovnostní podmínky a výsledky tohoto přístupu jsou následující:



Obrázek 4.16: Proložení fondů scénáři



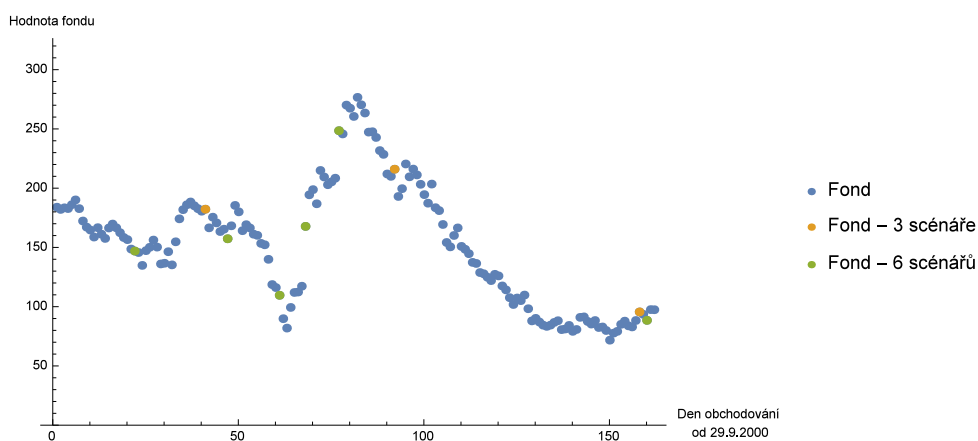
Obrázek 4.17: Proložení dluhopisů scénáři

### 4.4.3 Výběr z reálných dat

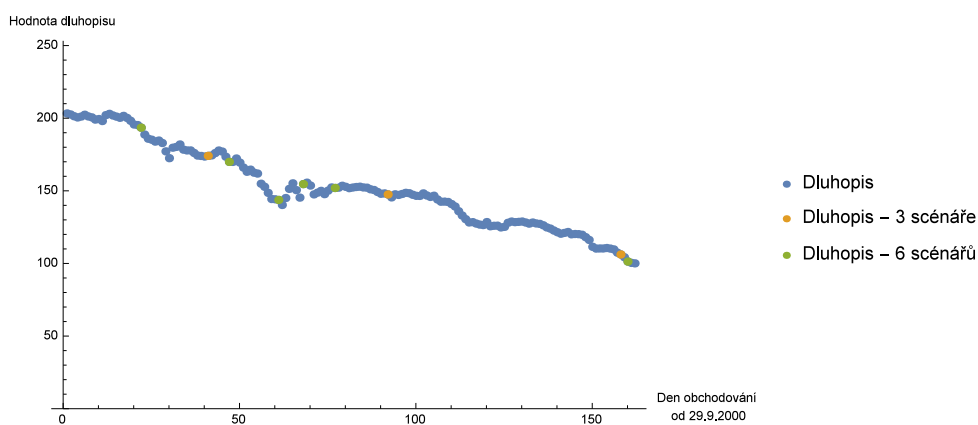
I úlohu na shodu prvních momentů a rozptylů je možné řešit pomocí výběru z reálných dat. Znamená to vyčíslit funkci

$$\begin{aligned} & \alpha_{1,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \right| + \\ & + \alpha_{2,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \right| + \\ & \alpha_{1,2} \left| \left( \frac{1}{S} \sum_{s=1}^S (\omega_1^s)^2 - \left( \frac{1}{S} \sum_{s=1}^S \omega_1^s \right)^2 \right) - \text{var}(1) \right| + \\ & + \alpha_{2,2} \left| \left( \frac{1}{S} \sum_{s=1}^S (\omega_2^s)^2 - \left( \frac{1}{S} \sum_{s=1}^S \omega_2^s \right)^2 \right) - \text{var}(2) \right| \end{aligned}$$

pro všechny  $S$ -tice a nalézt minimum a jemu odpovídající scénáře. Získáme tak následující scénáře:



Obrázek 4.18: Proložení fondů scénáři



Obrázek 4.19: Proložení dluhopisů scénáři

#### 4.4.4 Scénáře nalezené pro shodu prvních momentů a rozptylů

V následující tabulce (4.6) můžeme vidět, že všechny metody kromě výběru z reálných dat, přesně splnily požadavky. To znamená, že při generování nových scénářů pomocí minimalizace dané úlohy váhy  $\alpha_*$  opět nemají vliv na řešení. Opět mohou řešení ovlivnit pouze v případě výběru z reálných dat.

I.	II.	III.	IV.	V.	VI.	VII.
3	PR	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0
	M	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0
	G	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0
	RD	F	164,6	12,07	3586,66	1286,62
		D	142,58	-9,72	1170,76	389,67
		F+D		2,35		1676,29
6	PR	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0
	G	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0
	RD	F	141,34	-11,19	3671,05	1098,01
		D	152,5	0,3	937,337	156,25
		F+D		-10,89		1254,26
10	PR	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0
	G	F	152,53	0	2573,03	0
		D	152,3	0	781,09	0
		F+D		0		0

Tabulka 4.6: Popisné statistiky nalezených scénářů pro shodu rozptylů



Znak	Vysvětlivka
I.	$S$
II.	Způsob řešení úlohy
III.	Složka
IV.	$\frac{1}{S} \sum_{s=1}^S \omega^s$
V.	$\frac{1}{S} \sum_{s=1}^S \omega^s - \mu_1$
VI.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2 - (\frac{1}{S} \sum_{s=1}^S \omega^s)^2$
VII.	$\frac{1}{S} \sum_{s=1}^S (\omega^s)^2 - (\frac{1}{S} \sum_{s=1}^S \omega^s)^2 - \text{var}$
PR	Pravidelné a symetrické rozložení okolo střední hodnoty
M	Mathematica
G	GAMS
RD	Výběr z reálných dat
F	Fond
D	Dluhopis

Tabulka 4.7: Vysvětlivky

## 4.5 Shoda středních hodnot, rozptylů a korelací

Ještě si ilustrovme příklad, kdy bychom požadovali shodu nejen prvních momentů a rozptylů ale i korelací. V takovém případě je možné úlohu řešit pomocí nelineárního programování nebo metodou výběru z reálných dat.

Jako úloha nelineárního programování bude mít problém následující formulaci:

$$\begin{aligned}
& \min \alpha_{1,1}\gamma_{1,1} + \alpha_{2,1}\gamma_{2,1} + \alpha_{1,2}\gamma_{1,2} + \alpha_{2,2}\gamma_{2,2} + \beta_{1,2}\gamma \\
& \text{za podmínek } \gamma_{1,1} + \frac{1}{S} \sum_{s=1}^S \omega_1^s \geq \mu_1(1) \\
& \gamma_{1,1} - \frac{1}{S} \sum_{s=1}^S \omega_1^s \geq -\mu_1(1) \\
& \gamma_{2,1} + \frac{1}{S} \sum_{s=1}^S \omega_2^s \geq \mu_1(2) \\
& \gamma_{2,1} - \frac{1}{S} \sum_{s=1}^S \omega_2^s \geq -\mu_1(2) \\
& \gamma_{1,2} + \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \geq \text{var}(1) \\
& \gamma_{1,2} - \left( \sum_{s=1}^S \frac{1}{S} (\omega_1^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_1^s \right)^2 \right) \geq -\text{var}(1) \\
& \gamma_{2,2} + \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \geq \text{var}(2) \\
& \gamma_{2,2} - \left( \sum_{s=1}^S \frac{1}{S} (\omega_2^s)^2 - \left( \sum_{s=1}^S \frac{1}{S} \omega_2^s \right)^2 \right) \geq -\text{var}(2) \\
& \gamma + \frac{\sum_{s=1}^S (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s) (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s)}{\sqrt{\sum_{s=1}^S (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s)^2 \sum_{s=1}^S (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s)^2}} \geq \rho_{1,2} \\
& \gamma - \frac{\sum_{s=1}^S (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s) (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s)}{\sqrt{\sum_{s=1}^S (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s)^2 \sum_{s=1}^S (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s)^2}} \geq -\rho_{1,2} \\
& \omega_1^s \geq c_1 \\
& \omega_2^s \geq c_2.
\end{aligned}$$

Tuto úlohu bude z hlediska časové náročnosti (viz tabulka 4.8) vhodnější řešit pomocí softwaru GAMS než pomocí softwaru Mathematica.

Pokud bychom úlohu chtěli řešit pomocí metody výběru z reálných dat, zna-

menalo by to pro všechny  $S$ -tice vyčíslit funkci

$$\begin{aligned}
& \alpha_{1,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_1^s - \mu_1(1) \right| + \\
& + \alpha_{2,1} \left| \frac{1}{S} \sum_{s=1}^S \omega_2^s - \mu_1(2) \right| + \\
& + \alpha_{1,2} \left| \left( \frac{1}{S} \sum_{s=1}^S (\omega_1^s)^2 - \left( \frac{1}{S} \sum_{s=1}^S \omega_1^s \right)^2 \right) - \text{var}(1) \right| + \\
& + \alpha_{2,2} \left| \left( \frac{1}{S} \sum_{s=1}^S (\omega_2^s)^2 - \left( \frac{1}{S} \sum_{s=1}^S \omega_2^s \right)^2 \right) - \text{var}(2) \right| + \\
& + \beta_{1,2} \left| \rho_{1,2} - \frac{\sum_{s=1}^S (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s) (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s)}{\sqrt{\sum_{s=1}^S (\omega_1^s - \frac{1}{S} \sum_{s=1}^S \omega_1^s)^2 \sum_{s=1}^S (\omega_2^s - \frac{1}{S} \sum_{s=1}^S \omega_2^s)^2}} \right|
\end{aligned}$$

a nalézt minimum.

## 4.6 Časová náročnost výpočtů

Důležitou součástí výpočtů je bezesporu jejich časová náročnost. Některé výpočty, které dají přesnější výsledek, mohou trvat několikanásobně déle než jiné, méně přesné. V takovém případě je na zadavateli úlohy, aby se rozmyslel čemu dát přednost.

Ještě poznamenejme, že časy uvedené u řešení úloh řešených funkcí Minimize v softwaru Mathematica odpovídají součtu časů výpočtů pro  $i = 1, 2$ .

Úloha	Podmínky	Software	3 scénáře	6 scénářů	10 scénářů
4.2.1	I.	M	0.	0.	0.
		G	0.	0.	0.
	II.	M	0.	0.	0.
		G	0.	0.	0.
	III.	M	0.	0.	0.
		G	0.	0.	0.
4.2.2		C++	< min	< $\frac{1}{2}$ hod	
4.3.1	III.	M	< min	< min	< 5 min
		G	0.	0.	0.
4.3.2		C++	< min	< $\frac{1}{2}$ hod	
4.4.1		M	0.	0.	0.
4.4.2	III.	M	< min	< $\frac{1}{2}$ hod	
		G	< min	< min	< min
4.4.3		C++	< min	< hod	

Tabulka 4.8: Časová náročnost výpočtů

Znak	Vysvětlivka
I.	Okrajové podmínky pro první scénář
II.	Okrajové podmínky pro všechny scénáře
III.	Okrajové a nerovnostní podmínky pro všechny scénáře
M	Mathematica
G	GAMS

Tabulka 4.9: Vysvětlivky

## 4.7 Závěr

Z numerických výsledků můžeme vyvodit několik závěrů.

Prvním je, že generováním nových scénářů pomocí softwaru GAMS a Mathematica získáváme při řešení stejných úloh jiné scénáře. Všechny tyto výsledky jsou ale správné (lze nahlédnout v tabulkách 4.2, 4.4 a 4.6), protože dané úlohy obecně nemají jednoznačné řešení a záleží na jednotlivých algoritmech, které z více možných řešení vyberou.

Druhým z nich je, že, ač jsou úlohy na generování scénářů při podmínce shody středních hodnot časově výpočetně nenáročné, nemají příliš velký smysl, protože výsledky data nijak nereprezentují, jde pouze o několik bodů rozmístěných v blízkosti střední hodnoty. Naproti tomu scénáře získané výběrem z dat podstatně lépe data reprezentují, i když nemusí vždy dosáhnout přesné ani přibližné shody se střední hodnotou, avšak je to za cenu podstatně vyšší časové náročnosti.

Dalším závěrem je, že úlohy na shodu druhých momentů data reprezentují lépe než při shodě prvních momentů a i časová náročnost výpočtu je velmi dobrá. Již při řešení této úlohy je patrný rozdíl mezi rychlostí softwaru Mathematica a GAMS. První software je pro uživatele příznivější a méně náročný na porozumění. Druhý software počítá mnohokrát rychleji a veškerý čas vložený do náročnějšího implementování úlohy se vyplatí. Pro takovouto úlohu již výběr z reálných dat velmi ztrácí svou časovou náročností. Ač data reprezentuje vcelku dobře, protože se samovolně pohybují blíže střední hodnotě než v případě generování nových scénářů, není vhodné jej používat, protože počet  $n$ -tic, pro které musí vyčíslit funkci, roste příliš rychle se zvyšujícím se  $n$  a počtu dat.

Nejlépe na reprezentaci dat je generování nových scénářů při podmínce prvních momentů i rozptylů. Na obrázcích (4.14)-(4.17) můžeme vidět, že pokud bychom scénáře přeskládali do vhodnějšího pořadí, velmi dobře by reprezentovaly reálná data. Z hlediska časové náročnosti je nejvhodnější použít přístup generováním scénářů pomocí metody pravidelného a symetrického rozložení okolo střední hodnoty. Ovšem i metoda nelineárního programování v softwaru GAMS dobíhá za dobu kratší než jedna minuta, což je sice trochu déle než pomocí první metody, ale stále je to velmi dobrý čas. Naproti tomu výpočet pomocí softwaru Mathematica je z hlediska časové náročnosti velmi nevhodný, i když scénáře získané touto metodou jsou rovněž v pořádku.

Zároveň jsme mohli vypořádat, že při generování nových scénářů pomocí softwarů GAMS a Mathematica nemají váhy vliv na jednotlivá řešení uvedených příkladů, kdežto v případě výběru z reálných dat váhy vliv mohou mít, ale je to přímo závislé na tvaru dat.

# Literatura

- [1] ANDĚL, J. (2007). *Základy matematické statistiky*. Druhé vydání. Matfyzpress, Praha. ISBN 80-7378-001-1.
- [2] BUDÍKOVÁ, M., MIKOLÁŠ, Š. a OSECKÝ, P. (1996). *Teorie pravděpodobnosti a matematická statistika: sbírka příkladů*. Masarykova univerzita, Praha. ISBN 80-210-1329-X.
- [3] CONSEQ (2014). Data.
- [4] DUINTJER TEBBENS, E. J. (2012). *Analýza metod pro maticové výpočty: základní metody*. Matfyzpress, Praha. ISBN 978-80-7378-201-6.
- [5] DUPAČOVÁ, J., HURT, J. a ŠTĚPÁN, J. (2002). *Stochastic Modelling in Economics and Finance*. Applied optimization; 75. Kluwer Academic Publishers, Dodrecht. ISBN 1-4020-0840-6.
- [6] DUPAČOVÁ, J. A LACHOUT, P. (2011). *Úvod do optimalizace*. První vydání. Matfyzpress, Praha. ISBN 978-80-7378-176-7.
- [7] HAMALA, M. (1972). *Nelineární programovanie*. První vydání. Alfa, Bratislava.
- [8] HOYLAND, K., KAUT, M. a WALLACE, S. W. (2003). A heuristic for moment-matching scenarion generation. *Computational Optimization and Applications*, **24**, 169–185.
- [9] HOYLAND, K. A WALLACE, S. W. (2001). Generatiing scenario trees for multistage decision problems. *Management Science*, **47**(2), 295–307.
- [10] KOPÁČEK, J. (2003). *Matematická analýza pro fyziky II*. Matfyzpress, Praha. ISBN 80-86732-10-X.
- [11] KOUWENBERG, R. (2001). Scenario generation and stochastic programming models for asset liability management. *European Journal of Operational Research*, **134**, 279–292.
- [12] LUKŠAN, L. (1990). *Metody s proměnnou metrikou*. První vydání. Academia Praha, Praha. ISBN 80-200-0211-1.
- [13] LUKEŠ, J. A MALÝ, J. (2005). *Measure and Integral*. Druhé vydání. Matfyzpress, Praha. ISBN 80-86732-68-1.
- [14] PRÉKOPA, A. (1995). *Stochastic programming*. Akadémiai Kiadó, Budapest. ISBN 963-05-6872-1.

# Seznam obrázků

4.1	Použitá data . . . . .	23
4.2	Nevhodně spočítané 3 scénáře fondů . . . . .	24
4.3	Nevhodně spočítaných 6 scénářů dluhopisů . . . . .	24
4.4	Proložení fondů 3 scénáři . . . . .	25
4.5	Proložení dluhopisů 6 scénáři . . . . .	25
4.6	Proložení fondů scénáři . . . . .	26
4.7	Proložení dluhopisů scénáři . . . . .	26
4.8	Nalezené scénáře fondu . . . . .	27
4.9	Nalezené scénáře dluhopisu . . . . .	27
4.10	Nalezené scénáře fondu . . . . .	30
4.11	Nalezené scénáře dluhopisu . . . . .	30
4.12	Nalezené scénáře fondu . . . . .	31
4.13	Nalezené scénáře dluhopisu . . . . .	31
4.14	Proložení fondů scénáři . . . . .	33
4.15	Proložení dluhopisů scénáři . . . . .	33
4.16	Proložení fondů scénáři . . . . .	34
4.17	Proložení dluhopisů scénáři . . . . .	34
4.18	Proložení fondů scénáři . . . . .	35
4.19	Proložení dluhopisů scénáři . . . . .	35

# Seznam tabulek

4.1	Popisné statistiky sledovaných veličin . . . . .	23
4.2	Popisné statistiky nalezených scénářů pro shodu prvních momentů	28
4.3	Vysvětlivky . . . . .	29
4.4	Popisné statistiky nalezených scénářů pro shodu druhých momentů	32
4.5	Vysvětlivky . . . . .	32
4.6	Popisné statistiky nalezených scénářů pro shodu rozptylů . . . . .	36
4.7	Vysvětlivky . . . . .	37
4.8	Časová náročnost výpočtů . . . . .	39
4.9	Vysvětlivky . . . . .	40