

Cvičení z optimalizace Markowitzův model

Vojtěch Franc

5. prosince 2010

1 Úvod

Představme si, že máme určitou sumu peněz, kterou chceme zhodnotit investováním do akcií. V tomto případě stojíme před volbou, jakou část z naší celkové sumy peněz vložit do nákupu konkrétní akcie. Typicky se snažíme rozložit naši investici tak, abychom při co možná nejmenším riziku dosáhli požadovaného výnosu. S použitím Markowitzova modelu lze tuto úlohu vyjádřit jako problém kvadratického programování.

2 Neformální popis problému

Pro jednoduchost nejprve uvažujme, že chceme investovat do akcii pouze dvou ($n = 2$) hypotetických společností. Dále předpokládejme, že nás zajímá investice na dobu 1. měsíce. K rozhodnutí nám pomůže když si spočítáme, jaký by byl náš výnos, pokud bychom tuto investici udělali v předešlých měsících. Zelená křivka na obrázku 1 ukazuje příklad vývoje cen sledovaných v minulých $T = 100$ měsících. Z cen akcií si tedy můžeme spočítat výnosy v jednotlivých měsících

$$r_{i,t} = \frac{v_{i,t} - v_{i,t-1}}{v_{i,t-1}}, \quad t = 2, \dots, T, \quad i = 1, 2,$$

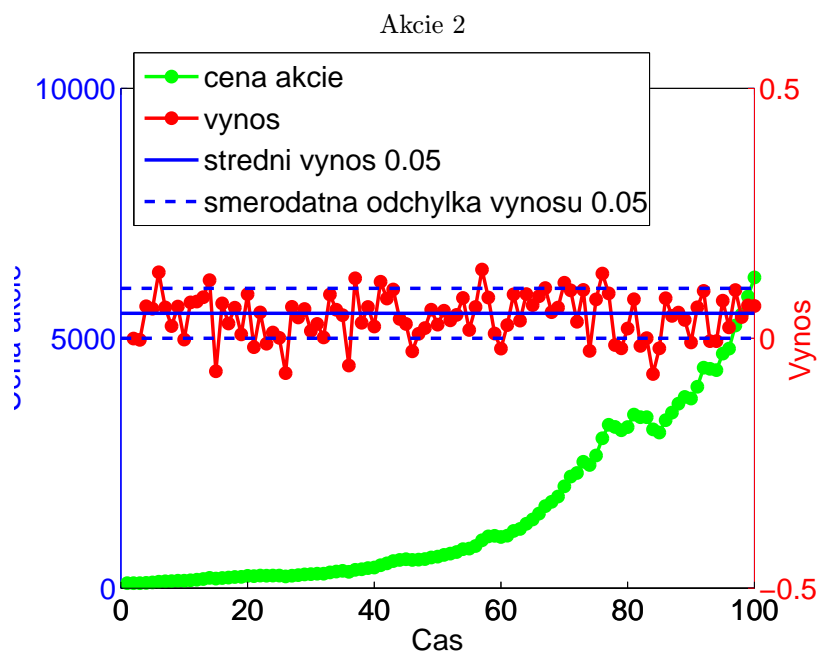
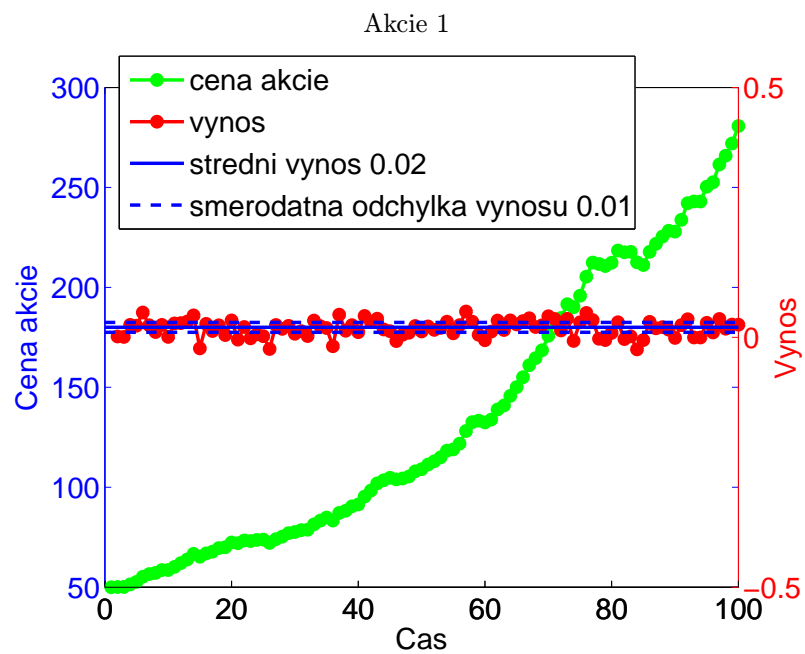
kde $v_{i,t}$ je cena i -té akcie v čase t a $r_{i,t}$ je výnos za období t . Výnosy jsou na obrázku 1 znázorněny červeně. Výnos $r_{i,t}$ udává poměr získaných a investovaných peněz pokud i -tou akcií v čase $t - 1$ nakoupíme a v čase t prodáme. Například nakoupíme-li akcií i v čase $t - 1$ za 1000 USD, pak prodejem v čase t získáme $1000r_{i,t}$ USD.

Dále se budeme zabývat klíčovou otázkou: Jak nám hodnoty minulých výnosů mohou pomoci při výběrů akcií, do které se vyplatí investovat v budoucnosti? Naše rozhodování opřeme o dvě statistiky, kterými jsou výběrový průměr a směrodatná odchylka. Výběrový průměr výnosů spočteme podle vzorce

$$\hat{r}_i = \frac{1}{T-1} \sum_{t=2}^T r_{i,t}, \quad i = 1, \dots, n.$$

Výběrový průměr je odhad průměrného, neboli očekávaného výnosu při dlouhodobém, tj. mnohokrát opakovaném, investování. Směrodatnou odchylku, respektive její druhou mocninu, spočteme podle vzorce

$$\hat{\sigma}_i^2 = \frac{1}{T-2} \sum_{t=2}^T (r_{i,t} - \hat{r}_i)^2, \quad i = 1, 2.$$



Obrázek 1: Vývoj ceny akcií 2. hypotetických společností a jejich výnos za 100 sledovaných období.

Směrodatná odchylka udává, jak moc se výnos odchyluje od své průměrné hodnoty. Předpokládejme pro jednoduchost, že výnosy jsou náhodné veličiny s normální rozdělením. V tomto případě bude výnos i -té akcie ležet v intervalu $\langle \hat{r}_i - \hat{\sigma}_i, \hat{r}_i + \hat{\sigma}_i \rangle$ se zhruba 68% pravděpodobností (tj. poměr počtu výnosů, které padnou do intervalu $\langle \hat{r}_i - \hat{\sigma}_i, \hat{r}_i + \hat{\sigma}_i \rangle$ ku počtu všech sledovaných výnosů, T , se bude s rostoucím T blížit číslu 0.68). Směrodatnou odchylku můžeme tedy vidět jako míru rizika, s jakým se skutečný výnos bude lišit od toho průměrného.

V našem příkladě z obrázku 1 je $\hat{r}_1 = 0.02$, $\hat{\sigma}_1 = 0.01$ a $\hat{r}_2 = 0.05$, $\hat{\sigma}_2 = 0.05$. To znamená, že s velkou pravděpodobností při investici do první akcie bude náš výnos 0.02 ± 0.01 (tj. $2 \pm 1\%$) a při investici do druhé akcie 0.05 ± 0.05 (tj. $5 \pm 5\%$). Tudíž do první akcie bude investovat ten, kdo upřednostňuje nižší výnos, jenž ale dosáhne s velkou jistotou. Naopak do druhé akcie budeme investovat ten, kdo chce velký výnos a nevadí mu vyšší nejistota. Je jasné, že v případě kdy jsou průměrné výnosy obou akcií stejné, budeme vždy investovat do akcie s nižší směrodatnou odchylkou.

Kromě nákupu jen jedné z akcií můžeme rozložit investovanou částku do obou současně. Zavedeme si váhy x_1 a x_2 , které budou udávat poměrnou částku investovanou do akcie 1 a 2. To znamená, že například z celkové částky 1000 USD investujeme $1000x_1$ USD do nákupu první akcie a $1000x_2$ USD do nákupu druhé akcie. Váhy nemohou být libovolná čísla, ale musí platit, že $x_1 \geq 0$, $x_2 \geq 0$ (uvažujeme jen nákup akcií) a současně $x_1 + x_2 = 1$ (obě investice dají dohromady celkovou částku). Pro tuto kombinovanou investici si můžeme opět spočítat průměrnou hodnotu

$$\hat{r}(x_1, x_2) = \frac{1}{T-1} \sum_{t=2}^T (r_{1,t}x_1 + r_{2,t}x_2) = \hat{r}_1x_1 + \hat{r}_2x_2 = \hat{\mathbf{r}}^T \mathbf{x},$$

a směrodatnou odchylku

$$\hat{\sigma}^2(x_1, x_2) = \frac{1}{T-2} \sum_{t=2}^T (r_{1,t}x_1 + r_{2,t}x_2 - \hat{r}(x_1, x_2))^2 = \mathbf{x}^T \hat{\mathbf{V}} \mathbf{x},$$

kde $\mathbf{x} = (x_1, x_2)^T$ je vektor vah, $\hat{\mathbf{r}} = (r_1, r_2)^T$ je vektor průměrů a $\hat{\mathbf{V}} \in \mathbb{R}^{2 \times 2}$ je kovarianční matice

$$\hat{V}_{i,j} = \frac{1}{T-2} \sum_{t=2}^T (r_{i,t} - \hat{r}_i)(r_{j,t} - \hat{r}_j), \quad i = 1, 2, j = 1, 2.$$

Všimněte si, že na diagonále kovarianční matice $\hat{\mathbf{V}}$ jsou kvadráty směrodatných odchylek $\hat{\sigma}_i^2$, $i = 1, 2$. Prvky mimo diagonálu udávají korelaci, neboli “souvztažnost” obou výnosů. Význam korelace můžeme intuitivně chápat takto: (i) pokud je korelace nulová, pak hodnoty výnosů z obou akcií na sobě nezávisí; (ii) při kladné hodnotě korelace bude s rostoucím výnosem z první akcie růst i výnos z té druhé; (iii) záporná hodnota korelace pak znamená, že růst výnosu z první akcie způsobí pokles výnosu z té druhé.

V našem příkladě, kdy počet akcií je pouze $n = 2$, můžeme volbou váhy x_1 v intervalu od 0 do 1 (váha $x_2 = 1 - x_1$) vyzkoušet všechny kombinace průměrného výnosu $\hat{r}(x_1, x_2)$ a směrodatné odchylky $\hat{\sigma}(x_1, x_2)$. Na základě těchto hodnot pak lze vybrat pro nás přijatelný kompromis mezi střední hodnotou výnosu a jeho směrodatnou odchylkou.

V obecném případě, kdy počet akcií je $n > 2$, může existovat nekonečně mnoho vah $\mathbf{x} \in \mathbb{R}^n$, při nichž bude hodnota průměrného výnosu stejná, ale směrodatná odchylka (riziko) se bude lišit. Výběr takové kombinace vah, která při daném minimálním výnosu zajistí minimální směrodatnou odchylku, lze vyjádřit jako úlohu kvadratického programování. Formální definice a řešení obecného případu jsou popsané v dalším odstavci.

3 Markowitzův model a úloha hledání optimálního portfolia

V tomto odstavci budeme používat následující značení:

n	počet akcií z nichž skládáme portfolio,
$\mathbf{x} = (x_1, \dots, x_n)^T$	vektor vah akcií v portfoliu, $\sum_{i=1}^n x_i = 1$, $x_i \geq 0$, $i \in \{1, \dots, n\}$,
$\boldsymbol{\rho} = (\rho_1, \dots, \rho_n)^T$	vektor náhodných proměnných odpovídající výnosům ve zvoleném období,
$\mathbf{r} = (r_1, \dots, r_n)^T$	vektor očekávaných výnosů ve zvoleném období,
r_p	minimální požadovaný výnos,
\mathbf{V}	kovarianční matice výnosů.

Jelikož budoucí výnosy z akcií nelze přesně určit, modelujeme je jako vektor náhodných veličin $\boldsymbol{\rho}$. Rozdělení pravděpodobnosti náhodného vektoru $\boldsymbol{\rho}$ lze charakterizovat prvním a druhým momentem, tj. vektorem středních hodnot $E(\boldsymbol{\rho}) = \mathbf{r}$ a kovarianční maticí $E((\boldsymbol{\rho} - \mathbf{r})(\boldsymbol{\rho} - \mathbf{r})^T) = \mathbf{V}$. Předpokládejme, že \mathbf{r} a \mathbf{V} jsou známé.

Při optimalizaci portfolia nás zajímají dvě kritéria: výnos a riziko. **Výnos portfolia** s vahami \mathbf{x} je definován jako střední hodnota celkového výnosu

$$r(\mathbf{x}) = E\left(\sum_{i=1}^n x_i \rho_i\right) = E(\mathbf{x}^T \boldsymbol{\rho}) = \mathbf{x}^T \mathbf{r}.$$

Riziko portfolia \mathbf{x} je definováno jako směrodatná odchylka celkového výnosu $\sigma(\mathbf{x})$, pro kterou platí

$$\sigma^2(\mathbf{x}) = E((\mathbf{x}^T \boldsymbol{\rho} - E(\mathbf{x}^T \boldsymbol{\rho}))^2) = E(\mathbf{x}^T (\boldsymbol{\rho} - \mathbf{r})(\boldsymbol{\rho} - \mathbf{r})^T \mathbf{x}) = \mathbf{x}^T E((\boldsymbol{\rho} - \mathbf{r})(\boldsymbol{\rho} - \mathbf{r})^T) \mathbf{x} = \mathbf{x}^T \mathbf{V} \mathbf{x}.$$

Naší úlohou je nalezení takového portfolia \mathbf{x}^* , tj. rozložení finančních prostředků do nákupu akcií, které zaručí, že výnos $r(\mathbf{x}^*)$ není menší než zadaná mez r_p a současně, že riziko $\sigma(\mathbf{x}^*)$ je minimální. Tuto úlohu lze vyjádřit jako problém kvadratického programování:

$$\mathbf{x}^* \in \underset{\mathbf{x} \in \mathbb{R}^n}{\operatorname{argmin}} \mathbf{x}^T \mathbf{V} \mathbf{x} \tag{1a}$$

za podmíněk

$$\sum_{i=1}^n x_i = 1, \tag{1b}$$

$$x_i \geq 0, \quad i = 1, \dots, n, \tag{1c}$$

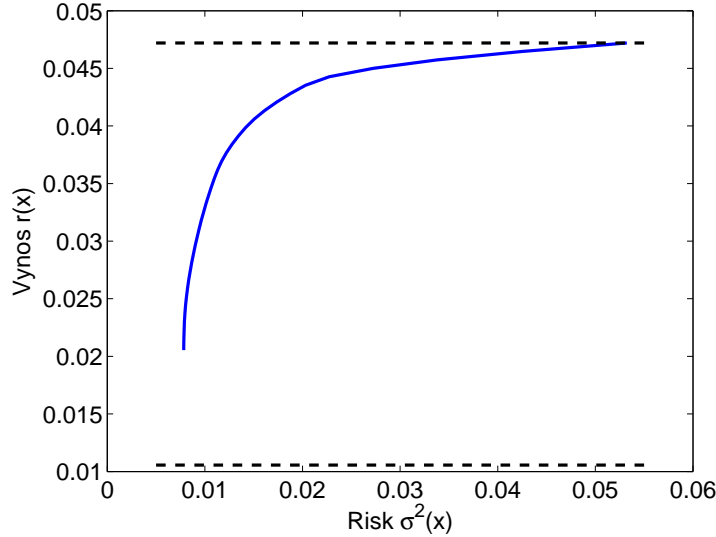
$$\sum_{i=1}^n x_i r_i \geq r_p. \tag{1d}$$

V Matlabu můžeme problém kvadratického programování řešit numericky pomocí funkce `quadprog`.

3.1 Odhad parametrů modelu

Markowitzův model předpokládá, že jak vektor středních hodnot \mathbf{r} , tak kovarianční matice \mathbf{V} jsou známé. Jak jsme si již ukázali v odstavci 2, lze tyto parametry odhadnout z dat, tj. ze známých časových řad vývoje akcií. Necht' $\{v_{i,t} \mid i = 1, \dots, n, t = 1, \dots, T\}$ jsou ceny akcií $i = 1, \dots, n$ v časech $t = 1, \dots, T$. Z cen akcií nejprve spočteme jejich výnosy

$$r_{i,t} = \frac{v_{i,t} - v_{i,t-1}}{v_{i,t-1}}, \quad i = 1, \dots, n, t = 2, \dots, T. \tag{2}$$



Obrázek 2: Závislost výnosu na riziku.

Ze vzorku výnosů $\{r_{i,t} \mid i = 1, \dots, n, t = 2, \dots, T\}$ pak můžeme odhadnout parametry (\mathbf{r}, \mathbf{V}) pomocí výběrového průměru a výběrové kovarianční matice

$$\hat{r}_i = \frac{1}{T-1} \sum_{t=2}^T r_{i,t}, \quad i = 1, \dots, n, \quad (3)$$

$$\hat{V}_{i,j} = \frac{1}{T-2} \sum_{t=2}^T (r_{i,t} - \hat{r}_i)(r_{j,t} - \hat{r}_j), \quad i = 1, \dots, n, j = 1, \dots, n. \quad (4)$$

V praxi pak předpokládáme, že se nedopustíme velké chyby pokud položíme $r_i = \hat{r}_i$, $i = 1, \dots, n$ a $V_{i,j} = \hat{V}_{i,j}$, $i = 1, \dots, n, j = 1, \dots, n$.

3.2 Data

V našem případě budeme plánovat investici do akcií 10. společností, které se obchodují na americké burze ¹. Konkrétně budeme uvažovat tyto společnosti: 1. Activision, 2. Amazon, 3. Apple, 4. Cisco, 5. Dell, 6. Human Genome, 7. Intel, 8. Microsoft, 9. Oracle a 10. Yahoo.

Budeme se zajímat o měsíční výnosy z akcií. Pro odhad parametrů modelu použijeme ceny akcií v období od ledna 1998 do července 2007 odečítané vždy začátkem měsíce. Všechny ceny jsou uvedené v USD a jsou v nich již započítané dividendy.

Účinnost Makowitzova modelu ověříme na vývoji akcií od srpna do prosince roku 2007. To znamená, že začátkem července nakoupíme akcie podle námi nalezené optimální strategie a budeme sledovat náš výnos ve zbytku roku.

3.3 Statistická interpretace

Pro úplnost ještě uvedeme statistickou interpretaci výnosu a rizika. Informace uvedená v tomto odstavci však není nezbytná k vypracování úlohy, ale pomůže ji lépe pochopit.

¹Data jsou stažena z <http://finance.yahoo.com/>

Výnos $r(\mathbf{x})$ a riziko $\sigma(\mathbf{x})$ popisují náhodnou veličinu $\rho = \mathbf{x}^T \boldsymbol{\rho}$, jejíž hodnota udává celkový výnos. Konkrétní realizaci náhodné veličiny ρ označíme písmenem r' . Za předpokladu, že ρ má normální (gaussovské) rozdělení, stačí znalost $r(\mathbf{x})$ a $\sigma(\mathbf{x})$ k jeho plnému popisu, tj.

$$\rho \approx N(r'; r(\mathbf{x}), \sigma^2(\mathbf{x})) = \frac{1}{\sigma(\mathbf{x})\sqrt{2\pi}} \exp\left(-\frac{(r' - r(\mathbf{x}))^2}{2\sigma^2(\mathbf{x})}\right).$$

Z rozdělení $N(r'; r(\mathbf{x}), \sigma^2(\mathbf{x}))$ lze spočítat pravděpodobnost náhodných jevů, které nás v této aplikaci zajímají. Například můžeme spočítat pravděpodobnost, že celkový výnos ρ bude menší než nějaká zvolená hodnota θ

$$P(\rho \leq \theta) = \int_{-\infty}^{\theta} N(r'; r(\mathbf{x}), \sigma^2(\mathbf{x})) dr'. \quad (5)$$

Hodnotu integrálu (5) lze v Matlabu spočítat pomocí funkce

```
P = cdf('Normal', theta, r, sigma^2)
```

4 Úkoly k vypracování

1. Stáhněte si soubor <http://cmp.felk.cvut.cz/cmp/courses/OPT/cviceni/10/akcie.mat>. Soubor obsahuje:

AkcieOdhad [115x10] Matice obsahující ceny (USD) akcii 10. společností sledovaných v období od ledna 1998 do července 2007, tj. za 115 měsíců. Tyto data použijte pro odhad parametrů Markowitzova modelu. Data jsou uložena tak, že **AkcieOdhad(1,:)** je vektor cen akcii z července 2007 a **AkcieOdhad(end,:)** jsou ceny z ledna 1998.

AkcieTest [5x10] Matice obsahující ceny (USD) akcii v období od srpna 2007 do prosince 2007. Tyto data použijte k otestování nalezeného rozložení portfolia. Data jsou uložena tak, že **AkcieTest(1,:)** jsou ceny akcii ze srpna 2007 a **AkcieTest(end,:)** ceny z prosince 2007.

Spolecnost {1x10 string} Jména společností v pořadí, které odpovídá uložení cen akcii v maticích **AkcieOdhad** a **AkcieTest**.

DatumOdhad [115x1 string] Data k nimž se vztahují ceny akcii použitých pro odhad modelu.

DatumTest [5x1 string] Data k nimž se vztahují ceny akcii použitých pro testování.

2. S použitím vzorců (2), (3) a (4) odhadněte parametry Markowitzova modelu z dat **AkcieOdhad**.
3. Interpretujte význam omezujících podmínek problému (1).
4. Jak by se změnila formulace problému (1), pokud bychom chtěli investovat do každé akcie maximálně 30% z celkové hodnoty portfolia ?
5. Vyřešte úlohu (1) pro r_p zadané v intervalu

```
r_p = [MinRp : (MaxRp-MinRp)/50 : MaxRp ]
```

kde $\text{MinRp} = \min(\mathbf{r})$, $\text{MaxRp} = \max(\mathbf{r})$ a \mathbf{r} je odhadnutý vektor středních hodnot. Pro takto spočtená rozložení portfolia \mathbf{x} vynesete do grafu hodnoty výnosu $r(\mathbf{x})$ a rizika $\sigma(\mathbf{x})$. Měli by jste získat křivku podobnou té na obrázku 2.

K řešení problému kvadratického programování použijte funkci `quadprog`.

6. Simulujte skutečný výnos portfolia získaného řešením úlohy (1) při $r_p = 0.037$. Uvažujte nákup akcii za cenu z července 2007, tj. za hodnoty `AkcieOdhad(1, :)`. Do grafu vynesete celkový výnos (vztažený k investici v červenci 2007) z cen akcii od srpna 2007 do prosince 2007, které jsou uloženy v matici `AkcieTest`.

Pro srovnání zobrazte výnosy získané při:

- (a) Rovnoměrném rozložení portfolia, tj. $x_i = \frac{1}{n}$, $i = 1, \dots, n$.
- (b) Investici do akcii s největším očekávaným, tj. $x_i = 0$, $i \neq j$ a $x_j = 1$, kde $j = \underset{i=1, \dots, n}{\operatorname{argmax}} r_i$.

7. Dokažte, že kovarianční matice \mathbf{V} je vždy pozitivně semidefinitní a tudíž, že problém (1) je konvexní.
8. Uvažujte situaci kdy kromě investice do n akcii máme navíc možnost investovat i do bezrizikového aktiva s výnosem r_0 . Např. můžeme naše peníze uložit do banky s garantovaným úrokem r_0 . Pro takto modifikovaný problém přeformulujte úlohu (1).

Reference

- [1] Lenka Slamová and Tereza Baumová. Markowitzův model. Studijní materiály MFF UK Praha, červen 2008.