Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Simona Macková

Diverzifikace v analýze obalu dat ve financích

Katedra pravděpodobnosti a matematické statistiky

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

Praha 2014

Velice ráda bych na tomto místě poděkovala panu RNDr. Martinu Brandovi, Ph.D. za odborné vedení, věcné připomínky a cenné rady poskytnuté k vypracování této bakalářské práce.

Prohlašuji, že jsem tuto bakalářskou práci vypracovala 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 22. 5. 2014 Simona Macková

Název práce: Diverzifikace v analýze obalu dat ve financích

Autor: Simona Macková

Katedra: Katedra pravděpodobnosti a matematické statistiky

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

Abstrakt: Práce pojednává o rozšířených modelech analýzy obalu dat a jejich aplikaci ve financích. Tato metoda umožňuje hodnotit efektivitu vybraných produkčních jednotek na základě vícero vstupů a výstupů. Při aplikaci ve financích můžeme využít jako vstupy například správní poplatky nebo míry rizika a jako výstupy očekávané výnosy sledovaných aktiv. Ukážeme základní klasické modely ve formě primárních i duálních úloh lineárního programování a později porovnáme s modely diverzifikačními. Pro využití ve financích a aplikaci na investiční aktiva je vhodné se diverzifikací, vzájemnými závislostmi aktiv, zabývat. Zde se již dostaneme do sféry nelineárního programování, a proto zavedeme vhodné míry vstupů a výstupů tak, aby úloha byla řešitelná. Především se zaměříme na podmíněnou hodnotu v riziku. Následně zavedeme model, který bude uvažovat diverzifikaci. Ten dále použijeme na reálná data vybraných podílových fondů.

Klíčová slova: analýza obalu dat, efektivita, diverzifikace, podmíněná hodnota v riziku

Title: Diversification in Data Envelopment Analysis in Finance

Author: Simona Macková

Department: Department of Probability and Mathematical Statistics

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

Abstract: This thesis deals with an extension of data envelopment analysis and its application in finance. This method enables to evaluate the efficiency of chosen production units based on several inputs and outputs. Administrative fees or risk measures can be used as inputs and expected incomes of observed assets as outputs in financial application. We show basic traditional models in a form of a primary problem of linear programming and a dual problem as well and later compare with diversification models. It is suitable to deal with diversification which enables to consider dependencies between assets in case of finance and investments. Than we get to nonlinear programming problem hence we introduce appropriate risk and return measures to make the problem solvable. Especially, we focus on the conditional value at risk. Next we introduce the model which deals with diversification. We use this on real data of chosen mutual funds.

Keywords: Data envelopment analysis, Efficiency, Diversification, Conditional value at risk

Obsah

Úvod

Již dlouhá léta je otázkou, jak nejlépe financovat a investovat, abychom maximalizovali výnosy a minimalizovali riziko. Touto problematikou se v 50. letech 20. století začal významně zabývat americký ekonom Harry Markowitz. Základy teorie portfolia uvedl v článku Markowitz (1952), na který navázala řada dalších. Ve svých dílech se zabývá investicemi do akciových portfolií. Uvědomil si, že k řešení bude vhodné využít optimalizační principy. V základních optimalizačních úlohách využívá jako míru rizika především rozptyl a očekávané výnosy jako výstupy.

K významnému pokroku a rozšířením přispívají pánové Charnes, Cooper a Rhodes ve svém díle ze 70. let Charnes a kol. (1978). Zde navrhují metodu analýzy obalu dat (*Data Envelopement Analysis*). Je to specializovaná modelová metoda pro hodnocení efektivity, výkonnosti a produktivity. Umožňuje testování produkčních jednotek za využití vícero vstupů a výstupů. Jedná se tedy o vícekriteriální rozhodování. Tím se daří metodu zobecnit a je aplikovatelná v mnoha odvětvích.

Modely analýzy obalu dat jsou především zajímavé pro aplikaci ve financích. Další články a publikace se zabývají využitím různých vstupů a výstupů relevantních pro investice a následnými dopady na výslednou efektivitu. Později se však klasické DEA modely ukazují jako matematicky slabší testy, neboť nezohledňují závislosti mezi jednotlivými aktivy v portfoliu. Jako řešení tohoto nedostatku se nabízelo využití diverzifikace. Modely zahrnující diverzifikaci se skutečně ukázaly jako silnější, neboť obecně hodnoty efektivnosti vycházely nižší. V textu Lamb a Tee (2012) je popsána jako vhodná míra rizika v diverzifikačním modelu ta, která je koherentní. Další rozšíření nabízí například článek Branda (2013).

Vývoj se v posledních 60. letech značně posunul – od Markowitzova klasického mean-risk modelu došlo ke zobecnění a rozšíření na mnohé modely, které ovšem stále nabízejí prostor k bádání.

V této práci se budeme zabývat diverzifikačním modelem a aplikací koherentních měr. V první kapitole stručně přiblížíme klasické modely analýzy obalu dat. Druhá kapitola popíše, jak zohlednit diverzifikaci do již popsaných modelů. Vhodnými mírami a jejich správným výpočtem se budeme zabývat v kapitole třetí a čtvrtá kapitola přinese propojení dosud nabytých poznatků a uvedení našeho výsledného modelu. Tímto modelem se budeme dále zabývat v páté kapitole, kde jej vyzkoušíme na reálných datech vybraných podílových fondů. V závěru zkusíme porovnat výsledky klasického DEA modelu a modelu diverzifikačního.

1 Analýza obalu dat

Analýza obalu dat neboli Data Envelopment Analysis (*DEA*) jsou optimalizační modely navržené pro hodnocení efektivnosti, popřípadě výkonnosti a produktivity homogenních produkčních jednotek. Homogenní produkční jednotky vykazují shodné nebo ekvivalentní výstupy a mají obdobné vstupy. Produkovanými výstupy jsou především kladné efekty, nějaké pozitivní výsledky. Každá jednotka spotřebovává určité vstupy a naším cílem je tyto vstupy minimalizovat. Zjednodušeně řečeno, pro výstupy platí čím vyšší, tím lepší a pro vstupy naopak, čím nižší, tím výhodnější.

1.1 Efektivnost

DEA modely se řadí mezi metody vícekriteriálního rozhodování a hodnocení efektivity jednotek. V případě jednoho vstupu a jednoho výstupu eficienci definujeme podílem

$$\frac{výstup}{vstup}.$$

Tento podíl v našem případě není až tak zajímavý, neboť nedozírnou výhodou DEA modelů oproti ostatním hodnotícím modelům je možnost využití více vstupů a více výstupů pro hodnocení.

DEA modely mají široké užití v různých odvětvích. Nás ovšem zajímá především aplikace ve financích. Později budeme uvažovat, že našimi produkčními jednotkami jsou různé podílové fondy, případně akcie. Mějme tedy n produkčních jednotek, kde každé má m vstupů a k výstupů. Vstupy označme x_{ij}, výstupy y_{lj} a dostaneme maticové vyjádření

$$X = \{x_{ij}; i = 1,2,\ldots,m; j = 1,2,\ldots,n\}$$
$$Y = \{y_{lj}; l = 1,2,\ldots,k; j = 1,2,\ldots,n\}.$$

Pak efektivnost q-té produkční jednotky zavádíme jako poměr vážených součtů

$$e_q = \frac{\sum_{l=1}^{k} \beta_l y_{lq}}{\sum_{i=1}^{m} \alpha_i x_{iq}}, \tag{1.1}$$

kde α_i a β_l jsou nezáporné váhy. Naším cílem bude určení těchto vah pomocí vhodného optimalizačního modelu.

Zvážíme-li všechny možné kombinace vstupů a výstupů dostáváme množinu přípustných možností. Hranice této množiny se nazývá efektivní hranice nebo obal dat. Přičemž platí, že ty jednotky, které jsou efektivní, na této hranici leží. V případě jednotek, které efektivní nejsou, DEA modely řeší, jak se vstupy či výstupy naložit, aby jednotky efektivní byly. Zavádí se následující modely:

1. Orientovaný na vstupy – navrhuje, jak snížit hodnotu vstupů
2. Orientovaný na výstupy – navrhuje, o kolik navýšit produkované výstupy
3. Kombinace obou – např. aditivní a odchylkové modely

1.2 Předpoklady modelu

Nyní se zaměříme na předpoklady modelu související s výnosy z rozsahu. Výnosy z rozsahu vystihují, jakým způsobem na sobě závisí proporcionální změny vstupů a následně vyvolané změny výstupů. Můžeme rozlišit konstantní a variabilní výnosy z rozsahu.

1.2.1 Konstantní výnosy z rozsahu

Konstantní výnosy z rozsahu neboli CRS (*Constant Returns of Scale*) jsou takové, pro něž platí, že pokud je jednotka (x,y) efektivní, pak je efektivní i $(x',y') = (\alpha x,\alpha y)$ pro $\alpha > 0$. Efektivní hranice je tedy kónický obal dat. Za efektivní jednotku považujeme tu, která na dané hranici leží, neboť nebude existovat žádná další jednotka, která by dosahovala vyšších výstupů při nižších nebo stejných vstupech.

Nabízí se využití relativní míry efektivnosti, kterou zavedeme jako

$$0 \leq \frac{e_q}{e} \leq 1, \tag{1.2}$$

kde e_q je efektivita q-té sledované jednotky a e je hodnota jednotky, kterou jsme za efektivní prohlásili. Tento podíl nám řekne, jak proporcionálně změnit vstupy či výstupy, aby sledovaná jednotka byla efektivní. Ať budeme model orientovat na vstupy nebo výstupy, výsledek podílu bude stejný, jak podrobněji vysvětluje Jablonský (2011).

1.2.2 Variabilní výnosy z rozsahu

V případě variabilních výnosů z rozsahu jinak VRS (*Variable Returns of Scale*), je hranice efektivnosti modifikována a vede na konvexní obal dat. Platí, že pokud považujeme za efektivní jednotky (x,y) a (x',y'), pak na efektivní hranici leží i jednotka $(\alpha x + (1-\alpha)x',\alpha y + (1-\alpha)y')$ pro $\alpha \in (0,1)$.

Relativní míra efektivnost (1.2) bude různá v závislosti na orientaci modelu. Proporcionální změna vstupů a výstupů vedoucí k efektivnosti nemůže být shodná. Pro variabilní výnosy z rozsahu bude stejně nebo více efektivních jednotek než v případě konstantních výnosů z rozsahu. Ovšem ty jednotky, které byly efektivní za předpokladu CRS, budou efektivní i pro VRS.

Porovnání obou podmínek a ukázky na jednoduchých příkladech jsou uvedeny v textu Jablonský (2011).

1.3 Vstupy

Jak již bylo zmíněno, možnost volby vstupů z celé řady různých kritérií je velikou výhodou DEA modelů. V případě námi uvažovaných podílových fondů by se daly zohlednit poplatky za správu či míry rizika. I v DEA modelech je naším cílem vstupy minimalizovat, respektive při dané hladině výstupů preferujeme menší hodnoty vstupů. Obvykle se užívá směrodatná odchylka, hodnota v riziku (*VaR = Value at Risk*) nebo podmíněná hodnota v riziku (*CVaR = Conditional Value at Risk*) na různých hladinách. V této práci je třeba se zaměřit především na diverzifikační modely, a proto, jak dále ukážeme, nás bude nejvíce zajímat CVaR. Konkrétním výpočtem a aplikací hodnot CVaR v modelu se budeme detailně zabývat v dalších kapitolách.

1.4 Výstupy

V případě, že budeme hodnotit jednotlivá aktiva, je nasnadě využít jako výstupy průměrný výnos – týdenní, měsíční, či roční. Průměrný čistý výnos q-té jednotky označíme $\bar{\rho}_q$. Ten však může nabývat i záporných hodnot, proto pokud bychom chtěli pracovat pouze s kladnými hodnotami, budeme uvažovat výnos hrubý

$$y_q = 1 + \bar{\rho}_q.$$

Na výstupy budeme pohlížet jako na očekávané výnosy.

1.5 Užívané modely

Nyní ukážeme nejzákladnější modely analýzy obalu dat. Lze tvrdit, že jednotlivé DEA modely se liší ve způsobu měření vzdálenosti jednotek v množině produkčních možností od efektivní hranice, tedy od obalu dat. Další rozšíření klasických modelů bychom nalezli v článku Branda a Kopa (2012).

1.5.1 CCR model

Model CCR dostal název podle svých tvůrců. Byl popsán pány Charnes, Cooper, Rhodes v textu Charnes a kol. (1978). Tento model využívá předpoklad CRS a snaží se maximalizovat míru efektivnosti konkrétní jednotky U_q jako podíl vážených výstupů a vážených vstupů (1.1) za předpokladu, že míry efektivnosti ostatních jednotek budou menší nebo rovny jedné. Pomocí vah získáme virtuální vstupy a výstupy. CCR modelem počítáme optimalizační úlohu, která nám určí váhy dat tak, aby míra efektivnosti sledovaného aktiva byla co nejpříznivější. Formulujme model pro jednotku U_q jako úlohu lineárního lomeného programování.

Maximalizace

$$\frac{\sum_{l=1}^{k} \beta_l y_{lq}}{\sum_{i=1}^{m} \alpha_i x_{iq}}, \tag{1.3}$$

za podmínky

$$\begin{aligned} \frac{\sum_{l=1}^{k} \beta_l y_{lj}}{\sum_{i=1}^{m} \alpha_i x_{ij}} &\leq 1, \qquad j = 1,2,\dots,n, \\ \alpha_i &\geq 0, \qquad i = 1,2,\dots,m, \\ \beta_l &\geq 0, \qquad k = 1,2,\dots,k. \end{aligned}$$

Úlohu bychom chtěli převést na standardní lineární programování. Použijeme Charnes-Cooperovu transformaci. Tato transformace je blíže uvedená v textu Charnes a kol. (1978) a převádí úlohu lineárního lomeného programování na úlohu programování lineárního. Tím dostáváme následující optimalizační úlohu.

Maximalizace

$$\sum_{l=1}^{k} \beta_l y_{lq}, \tag{1.4}$$

za podmínky

$$\begin{aligned}
\sum_{l=1}^{k} \beta_l y_{lj} &\leq \sum_{i=1}^{m} \alpha_i x_{ij}, \qquad j = 1,2,\ldots,n,\\
\sum_{i=1}^{m} \alpha_i x_{iq} &= 1,\\
\alpha_i &\geq 0, \qquad i = 1,2,\ldots,m,\\
\beta_l &\geq 0, \qquad l = 1,2,\ldots,k.
\end{aligned}$$

Tento tvar reprezentuje primární CCR model orientovaný na vstupy. Pro výpočet nás ovšem bude zajímat duální tvar úlohy. Jak se odvozuje tvar duální, ukážeme později u dalšího modelu (1.7). V tomto případě úloha vypadá následovně.

Minimalizace

$$\Theta_q, \tag{1.5}$$

za podmínky

$$\begin{aligned}
\sum_{j=1}^{n} \lambda_j y_{lj} &\geq y_{lq}, \qquad l = 1,2,\ldots,k,\\
\sum_{j=1}^{n} \lambda_j x_{ij} &\leq \Theta_q x_{iq}, \qquad i = 1,2,\ldots,m,\\
\lambda_j &\geq 0, \qquad j = 1,2,\ldots,n.
\end{aligned}$$

Proměnnými tohoto duálního modelu je vektor vah $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_n)$ a Θ_q. Proměnná Θ_q vyjadřuje míru efektivnosti sledované jednotky U_q, bude nebývat hodnot větších nebo rovných nule a menších nebo rovných jedné. Lze ji interpretovat jako míru redukce vstupů potřebnou k tomu, aby jednotka dosáhla efektivní hranice.

V některých případech bychom mohli požadovat, aby váhy v primárním modelu α_i a β_l nebyly pouze nezáporné, ale aby byly větší než nějaká infinitezimální konstanta ε, tedy nekonečně malé číslo. Takto bychom zaručili pro všechny jednotky kladné váhy. Jednoduše řečeno chceme, aby každá jednotka byla alespoň trochu zastoupena. Přidáním podmínek $\alpha_i \geq \varepsilon$, $\beta_l \geq \varepsilon$ dostáváme i rozdílnou duální úlohu. V tomto případě se nám přidají nové vektory skluzových proměnných s^+ a s^- a s úsporným maticovým zápisem X, Y dostáváme následující úlohu, k jejímuž postupnému odvození se dostaneme při demonstraci dalšího modelu.

Minimalizace

$$\Theta_q - \varepsilon(e^T s_i^+ + e^T s_i^-), \tag{1.6}$$

za podmínky

$$\begin{aligned} X\lambda + s^- &= \Theta_q x_q, \\ Y\lambda - s^+ &= y_q, \\ \lambda, s^+, s^- &\geq 0. \end{aligned}$$

Zde představuje vektor $e^T = (1,1,\ldots,1)$, x_q a y_q jsou vektory vstupů a výstupů jednotky U_q. Jednotka je efektivní, pokud pro optimální vektor vah λ^* platí $X\lambda^* = x_q$ a $Y\lambda^* = y_q$. To nastává v případě, že optimální skluzové proměnné s^+, s^- jsou rovny nule a $\Theta_q = 1$. Model orientovaný na výstupy bychom sestavili analogicky.

1.5.2 BCC model

Model BCC byl navržen v článku Banker a kol. (1984) pány Bankerem, Charnesem a Cooperem. Jeho předpokladem jsou variabilní výnosy z rozsahu, proto obal dat bude konvexní. Tento předpoklad přináší, že za efektivní bude označen vyšší nebo stejný počet jednotek. Výsledné optimalizační úlohy dosáhneme obdobnými kroky jako v předchozím modelu CCR, proto zde uvedeme rovnou model po aplikaci Charnes-Cooperovy transformace s předpoklady $\alpha_i \geq \varepsilon$, $\beta_l \geq \varepsilon$.

Maximalizace

$$\sum_{l=1}^{k} \beta_l y_{lq} + \mu, \tag{1.7}$$

za podmínky

$$\begin{aligned} \sum_{l=1}^{k} \beta_l y_{lj} + \mu &\leq \sum_{i=1}^{m} \alpha_i x_{ij}, \qquad j = 1,2,\ldots,n, \\ \sum_{i=1}^{m} \alpha_i x_{iq} &= 1, \\ \alpha_i &\geq \varepsilon, \qquad i = 1,2,\ldots,m, \\ \beta_l &\geq \varepsilon, \qquad l = 1,2,\ldots,k, \\ \mu &\in \mathbb{R}. \end{aligned}$$

Jednotlivé modely se od sebe liší v tom, jakých hodnot může nabývat proměnná μ. V modelu BCC to mohou být libovolné hodnoty, jak záporné, kladné i nulové. V dříve uvedeném modelu CCR se tato proměnná nevyskytla, neboť se jedná o odchylku od konstantního výnosu z rozsahu.

Jistě stojí za to ukázat, jakým způsobem přejdeme k duální formě úlohy a jak se ve výsledku projeví konvexita způsobená předpokladem variabilních výnosů z rozsahu.

		β_1	β_2	$\dots$	β_k	μ	α_1	α_2	$\dots$	α_m		
		$\geq$	$\geq$	$\dots$	$\geq$	$\in \mathbb{R}$	$\geq$	$\geq$	$\dots$	$\geq$		
		y_{11}	y_{21}	$\dots$	y_{k1}	1	$-x_{11}$	$-x_{21}$	$\dots$	$-x_{m1}$	$\leq$	0
		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
		y_{1n}	y_{2n}	$\dots$	y_{kn}	1	$-x_{1n}$	$-x_{2n}$	$\dots$	$-x_{mn}$	$\leq$	0
		0	0	$\dots$	0	0	x_{1q}	x_{2q}	$\dots$	x_{mq}	$=$	1
		1	0	$\dots$	0	0	0	0	$\dots$	0	$\geq$	ε
		0	1	$\dots$	0	0	0	0	$\dots$	0	$\geq$	ε
		$\vdots$		$\ddots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
		0	0	$\dots$	1	0	0	0	$\dots$	0	$\geq$	ε
		0	0	$\dots$	0	0	1	0	$\dots$	0	$\geq$	ε
		0	0	$\dots$	0	0	0	1	$\dots$	0	$\geq$	ε
		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\ddots$		$\vdots$	$\vdots$
		0	0	$\dots$	0	0	0	0	$\dots$	1	$\geq$	ε
		y_{1q}	y_{2q}	$\dots$	y_{kq}	1	0	0	$\dots$	0	max	

Tabulka 1.1: Primární model rozepsaný do tabulkové podoby

		β_1	β_2	$\dots$	β_k	μ	α_1	α_2	$\dots$	α_m		
		$\geq$	$\geq$	$\dots$	$\geq$	$\in \mathbb{R}$	$\geq$	$\geq$	$\dots$	$\geq$		
λ_1	$\geq$	y_{11}	y_{21}	$\dots$	y_{k1}	1	$-x_{11}$	$-x_{21}$	$\dots$	$-x_{m1}$	$\leq$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
λ_n	$\geq$	y_{1n}	y_{2n}	$\dots$	y_{kn}	1	$-x_{1n}$	$-x_{2n}$	$\dots$	$-x_{mn}$	$\leq$	0
Θ_q	$\in \mathbb{R}$	0	0	$\dots$	0	0	x_{1q}	x_{2q}	$\dots$	x_{mq}	$=$	1
$-s_1^+$	$\leq$	1	0	$\dots$	0	0	0	0	$\dots$	0	$\geq$	ε
$-s_2^+$	$\leq$	0	1	$\dots$	0	0	0	0	$\dots$	0	$\geq$	ε
$\vdots$	$\vdots$	$\vdots$		$\ddots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
$-s_r^+$	$\leq$	0	0	$\dots$	1	0	0	0	$\dots$	0	$\geq$	ε
$-s_1^-$	$\leq$	0	0	$\dots$	0	0	1	0	$\dots$	0	$\geq$	ε
$-s_2^-$	$\leq$	0	0	$\dots$	0	0	0	1	$\dots$	0	$\geq$	ε
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\ddots$		$\vdots$	$\vdots$
$-s_m^-$	$\leq$	0	0	$\dots$	0	0	0	0	$\dots$	1	$\geq$	ε
		$\geq$	$\geq$	$\dots$	$\geq$	$=$	$\geq$	$\geq$	$\dots$	$\geq$		min
		y_{1q}	y_{2q}	$\dots$	y_{kq}	1	0	0	$\dots$	0	max	

Tabulka 1.2: Doplnění tabulky pro odvození duálního modelu

Tvrzení 1. *Duální úloha k úloze primární* (1.7) *má tvar:*
Minimalizace

$$\Theta_q - \varepsilon(e^T s^+ + e^T s^-), \tag{1.8}$$

za podmínky

$$\begin{aligned} X\lambda + s^- &= \Theta_q x_q, \\ Y\lambda - s^+ &= y_q, \\ e^T\lambda &= 1, \\ \lambda, s^+, s^- &\geq 0. \end{aligned}$$

Důkaz. Nejprve úlohu přepíšeme do tabulky 1.1. Do horního řádku uvedeme proměnné primárního modelu a dále po řádcích zapisujeme podmínky. Ve spodním řádku vyplníme koeficienty účelové funkce. Pro přechod k duálnímu problému příslušně doplníme řádky i sloupce, přičemž budeme postupovat způsobem uvedeným v textu o základech optimalizace Dupačová a Lachout (2011). Tvar modelu (1.8) získáme náležitým přepisem tabulky 1.2. V prvním sloupci jsou zapsány proměnné duálního tvaru a druhý sloupec hovoří o intervalech, na kterých jsou definované. Sloupec úplně vpravo udává koeficienty účelové funkce a poslední vyplněné políčko říká, že se bude jednat o minimalizaci této funkce. Omezení duálního modelu zjistíme čtením středové části po sloupcích.

❑

V poslední kapitole této práce budeme model aplikovat na konkrétní data a použijeme model BCC v podobě (1.9). Jak již dříve bylo naznačeno u modelu CCR (1.5), tento duální tvar získáme, pokud váhy v primárním modelu splňují $\alpha_i \geq 0$, $\beta_l \geq 0$.

Minimalizace

$$\Theta_q, \tag{1.9}$$

za podmínky

$$\begin{aligned} \sum_{j=1}^{n} \lambda_j y_{lj} &\geq y_{lq}, & l &= 1,2,\dots,k, \\ \sum_{j=1}^{n} \lambda_j x_{ij} &\leq \Theta_q x_{iq}, & i &= 1,2,\dots,m, \\ \sum_{j=1}^{n} \lambda_j &= 1, \\ \lambda_j &\geq 0, & j &= 1,2,\dots,n. \end{aligned}$$

2 Diverzifikace v analýze obalu dat

České slovo pro diverzifikaci bychom mohli hledat ve výrazu rozrůznění. Popřípadě bychom ho mohli interpretovat jako strategii, která nespoléhá na jediný produkt, ale zájmy rozděluje do různých oblastí, aby snížila riziko. Zejména v oblasti financí je zajímavé uvažovat závislost mezi jednotlivými aktivy. V případě analýzy obalu dat říkáme, že diverzifikace je jakási možnost, že kombinace aktiv bude mít nižší riziko nebo vyšší výnos než průměrné riziko a výnosy jednotlivých samostatných aktiv. Diverzifikace je úzce spjata s korelací. Pokud jsou aktiva dokonale kladně korelovaná, pro diverzifikaci již není prostor.

Ukážeme si matematickou úlohu, která diverzifikaci zohledňuje. Taková úloha ovšem nemusí být prakticky řešitelná. Převedeme ji proto na úlohu konvexního programování. Konvexitu úlohy nám zaručí předpoklad koherentních měr rizika, jako je například podmíněná hodnota v riziku.

2.1 Obecný diverzifikační model

Pro další uváděné modely se budeme co nejvíce držet již zavedeného značení a již odvozeného duálního modelu orientovaného na vstupy s předpoklady VRS. V klasickém DEA modelu pracujeme s maticemi vstupů X a výstupů Y, v nichž jsou zahrnuta data, která se zdají být vhodná. Tentokrát se blíže zaměříme na finance, především akcie a podílové fondy. Označíme R_q náhodný vektor výnosů q-tého aktiva. Písmenem F_i označme i-tou míru vstupů, konkrétněji míru rizika. Písmenem G_l označme l-tou míru výnosů, tedy výstupy. Obecně nás bude zajímat úloha (2.1) analogie modelu uvedeného v článku Lamb a Tee (2012).

Minimalizace

$$\Theta_q, \tag{2.1}$$

za podmínky

$$\begin{aligned}
G_l\left(\sum_{j=1}^{n}\lambda_j R_j\right) &\geq G_l(R_q), \qquad l = 1,2,\ldots,k,\\
F_i\left(\sum_{j=1}^{n}\lambda_j R_j\right) &\leq \Theta_q F_i(R_q), \qquad i = 1,2,\ldots,m,\\
\sum_{j=1}^{n}\lambda_j &= 1,\\
\lambda_j &\geq 0, \qquad j = 1,2,\ldots,n.
\end{aligned}$$

Toto je diverzifikační model orientovaný na vstupy s předpokladem VRS – variabilních výnosů z rozsahu. Každý fond je zde porovnáván se všemi možnými portfolii, která jsou z daných aktiv sestavitelná, a tak model zohledňuje diverzifikaci. Přesněji řečeno vstupy a výstupy každého fondu jsou porovnávány se vstupy a výstupy všech možných portfolií. Zásadním problémem tohoto modelu je, že ne vždy bude početně řešitelný, neboť se jedná o obtížnou nelineární úlohu.

Bylo již publikováno několik různých studií, které se zabývají diverzifikací, ale ne ve formě plných diverzifikačních modelů. Například byl popsán diverzifikační lineární DEA model, který zvažuje právě jeden vstup a jeden výstup Lozano a Gutiérrez (2008). Článek Lamb a Tee (2012) zase navrhuje iterační početní proces. V této práci navrhneme převedení obecného diverzifikačního modelu na problém konvexního programování.

2.2 Převod na konvexní problém

Definice 1. *Konvexní programování je obecně reprezentováno optimalizační úlohou*

$$\min_{x\in M} f(x),$$

přičemž $f(x)$ je konvexní funkce, množina M je konvexní, neboť je popsána soustavou nelineárních nerovnic $g_i(x) \leq 0$, kde funkce $g_i(x)$ pro $i = 1,\ldots,m$, jsou taktéž konvexní.

Jak již bylo uvedeno, diverzifikační model bude lépe řešitelný, pokud se nám podaří nalézt vhodné míry vstupů a výstupů tak, aby se problém stal konvexním. Bude třeba, aby množina produkčních možností splnila předpoklad konvexity. Tu zaručíme variabilními výnosy z rozsahu, tedy podmínkou VRS. Algoritmy obvykle zjišťují lokální řešení, konvexita nám zajistí, že toto řešení bude i globální.

Označme $L = \left\{\sum_{j=1}^{n}\lambda_j R_j : \sum_{j=1}^{n}\lambda_j = 1, \lambda_j \geq 0, j = 1,\ldots,n\right\}$. L bychom mohli interpretovat jako množinu všech dosažitelných výnosů portfolia. Pro vstupy budeme hledat vhodnou míru, která splní koherentní vlastnosti. Hledáme míru rizika $F : L \to \mathbb{R}$. Tu zadefinujeme obdobně jako Lamb a Tee (2012).

Definice 2. *Řekněme, že míra rizika F je koherentní, jestliže splňuje vlastnosti:*

$$\begin{aligned}
\textit{Ekvivariance vůči posunutí:} &\quad \textit{pro } l \in L \textit{ a } \alpha \in \mathbb{R} : F\ (l + \alpha r) = F(l) - \alpha \\
\textit{Subaditivita:} &\quad \textit{pro } l_1, l_2 \in L : F(l_1 + l_2) \leq F(l_1) + F(l_2) \\
\textit{Pozitivní homogenita:} &\quad \textit{pro } \lambda \geq 0 \textit{ a } l \in L : F(\lambda l) = \lambda F(l) \\
\textit{Monotonie:} &\quad \textit{pro } l_1, l_2 \in L, \textit{ kdy platí } l_1 \leq l_2 : F(l_1) \geq F(l_2).
\end{aligned}$$

Subaditivita zaručuje, že součet náhodných veličin nebude mít větší riziko než součet rizik jednotlivých samostatných náhodných veličin. Pozitivní homogenita nám říká, že riziko je úměrné výši investovaných prostředků. Tyto dvě vlastnosti implikují naši požadovanou konvexitu. Užívanou mírou rizika splňující podmínky koherentní míry je podmíněná hodnota v riziku (*CVaR*). Této míře se budeme podrobně věnovat ve třetí kapitole.

Nyní ukážeme vlastnosti, které budeme požadovat pro míru výnosů $G : L \to \mathbb{R}$. Tuto míru zavedeme jako zápornou koherentní, tedy platí $F = -G$.

Definice 3. *Řekněme, že míra G je mírou výnosů, jestliže splňuje vlastnosti:*

$$\begin{aligned}
\textit{Ekvivariance vůči posunutí:} &\quad \textit{pro } l \in L \textit{ a } \alpha \in \mathbb{R} : G\ (l + \alpha) = G(l) + \alpha \\
\textit{Superaditivita:} &\quad \textit{pro } l_1, l_2 \in L : G(l_1 + l_2) \geq G(l_1) + G(l_2) \\
\textit{Pozitivní homogenita:} &\quad \textit{pro } \lambda \geq 0 \textit{ a } l \in L : G(\lambda l) = \lambda G(l) \\
\textit{Monotonie:} &\quad \textit{pro } l_1, l_2 \in L, \textit{ kdy platí } l_1 \leq l_2 : G(l_1) \leq G(l_2).
\end{aligned}$$

Takové vlastnosti měr vstupů a výstupů nám zaručí konvexitu modelu.

Tvrzení 2. *Předpokládejme, že vstupy v diverzifikačním modelu jsou koherentní míry a výstupy míry výnosů. Pak je diverzifikační model konvexní úloha.*

Důkaz. Toto tvrzení lze dokázat přímo z definice konvexity a vlastností měr F a G, konkrétně subaditivity, superaditivity a pozitivní homogenity. Důkaz je podrobně popsán v článku Lamb a Tee (2012).

❑

Koherentní míry rizika budou důležité pro splnění konvexity. Všechny požadované vlastnosti splňuje již výše zmiňovaná míra CVaR_α. Za hladinu α budeme dosazovat různé hodnoty, například $0.75, 0.95$. Tím získáme vícero měr, které dále užijeme v modelu.

Koherentní míry rizika ovšem mohou nabývat i záporných hodnot. My bychom rádi pro náš model získávali hodnoty pouze nezáporné. Pokud je míra rizika záporná, potom investice bude zaručeně výnosná. Pro výpočet bude vhodné tyto záporné hodnoty nahradit nulou.

Tvrzení 3. *Mějme množinu náhodných veličin L. Předpokládejme koherentní míru rizika $F : L \to \mathbb{R}$. Potom $\hat{F} = max(F,0)$ splňuje pozitivní homogenitu, subaditivitu a monotonii.*

Je zřejmé, že tato míra $\hat{F}$ nemusí vždy splňovat vlastnost ekvivariance vůči posunutí. To pro náš model nebude až tak vadit, neboť jedinou výhodou této vlastnosti je, že bychom mohli započíst správní poplatky podílových fondů.

Pokud míra rizika CVaR_α vyhovuje konvexitě, pak určitě bude $-\text{CVaR}_\alpha$ vyhovovat požadavkům jako míra výnosu. Navrhněme tedy míru rizika tak jako v textu Lamb a Tee (2012) CVaR_α pro $\alpha \geq 0.6$ a $-\text{CVaR}_\alpha$ pro $\alpha \leq 0.4$ jako míru výnosu.

3 Míry rizika

Způsobů měření rizika je mnoho a v klasických DEA modelech je pouze na nás, které z nich považujeme za relevantní a vybereme do našich výpočtů. Kromě některých méně častých metod se obvykle užívá směrodatná odchylka, VaR_α a CVaR_α. Bylo naznačeno, že pro diverzifikační model bude CVaR_α ideální mírou rizika. Detailně se tímto tématem zabývá článek Rockafellar a Uryasev (2000), zjednodušené pojetí nabízí Sůva (2009).

3.1 Hodnota v riziku

Hodnota v riziku na hladině α (VaR_α) má své nevýhody. Nesplňuje podmínku subaditivity, a tak se neřadí ke koherentním mírám rizika. Ve druhé kapitole jsme zmínili, že právě tyto vlastnosti nám budou vyhovovat. Dále VaR_α nenese žádnou informaci o tom, jak závažné mohou být ztráty, pokud by hodnota v riziku byla překročena, tedy v případech nastávajících ve $(1-\alpha).100\%$. Vylepšení nám zajistí právě CVaR_α. Uvedeme definici VaR_α pro následnou lepší orientaci.

Máme-li náhodné výnosy R, potom označme ztrátu $Z = -R$. Distribuční funkcí ztráty jako náhodné veličiny je

$$\Psi(\zeta) = P[Z \leq \zeta].$$

Definice 4. *Mějme ztrátu označenou jako Z a $\alpha \in (0,1)$, pak hodnotu v riziku VaR_α definujeme jako*

$$VaR_\alpha = \min_{\zeta} \Psi(\zeta) \geq \alpha.$$

VaR_α je α-kvantilem funkce náhodné veličiny ztráty Z.

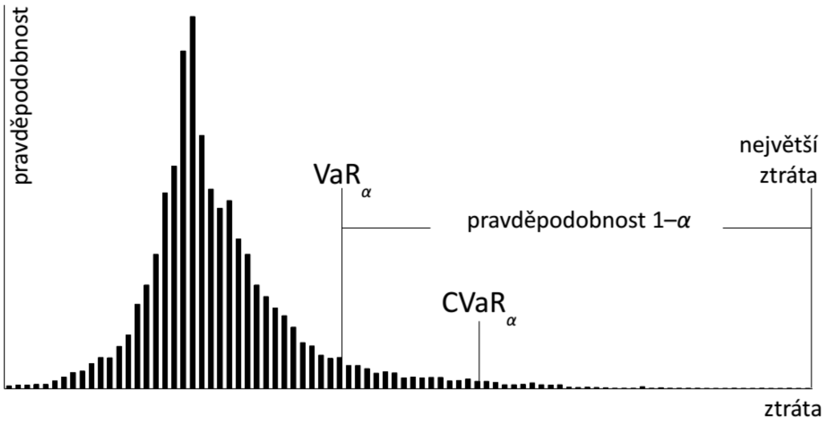
3.2 Podmíněná hodnota v riziku

CVaR_α podmíněnou hodnotu v riziku lze interpretovat jako střední hodnota α-chvostu rozdělení náhodné veličiny Z.

Definice 5. *Definujeme náhodnou veličinu T_α, kterou nazveme α-chvost náhodné veličiny a jejíž distribuční funkce odpovídá*

$$\Psi_\alpha(\zeta) = \begin{cases} 0, & \textit{pro} \quad \zeta < VaR_\alpha, \\ \dfrac{\Psi(\zeta) - \alpha}{1 - \alpha}, & \textit{pro} \quad \zeta \geq VaR_\alpha. \end{cases}$$

Definujeme potom $CVaR_\alpha = E\, T_\alpha$.



Obrázek 3.1: VaR_α a CVaR_α v grafu

Obrázek 3.1 převzatý z textu Sůva (2009) ilustruje, jak veličiny VaR_α a CVaR_α rozdělují graf hustoty.

Následující tvrzení bylo převzato z textu Rockafellar a Uryasev (2002).

Tvrzení 4. *Mějme náhodný vektor výnosů R a vektor ztrát $Z = -R$ s diskrétním rozdělením. Seřaďme ztráty Z podle velikosti $Z^{[1]} < Z^{[2]} < \cdots < Z^{[S]}$ s pravděpodobností $p_s > 0$ pro $Z^{[s]}$. Nechť index s_α splňuje*

$$\sum_{s=1}^{s_\alpha-1} p_s < \alpha \leq \sum_{s=1}^{s_\alpha} p_s,$$

potom je $CVaR_\alpha$ dáno funkcí

$$CVaR_\alpha = \frac{1}{1-\alpha}\left[\left(\sum_{s=1}^{s_\alpha} p_s - \alpha\right) Z^{[s_\alpha]} + \sum_{s=s_\alpha+1}^{S} p_s Z^{[s]}\right]. \tag{3.1}$$

Tento vzorec můžeme usnadnit tím, že budeme předpokládat rovnoměrné rozdělení výnosů. Tedy každý výnos je stejně pravděpodobný a v případě S výnosů nastane s pravděpodobností $\frac{1}{S}$, potom lze psát

$$\frac{s_\alpha - 1}{S} < \alpha \leq \frac{s_\alpha}{S}$$

a pro podmíněnou hodnotu v riziku platí

$$CVaR_\alpha = \frac{1}{1-\alpha}\left[\left(\frac{s_\alpha}{S} - \alpha\right) Z^{[s_\alpha]} + \frac{1}{S}\sum_{s=s_\alpha+1}^{S} Z^{[s]}\right]. \tag{3.2}$$

Dále uvažujme vektor λ, kterým rozumíme váhy pozorovaného portfolia. Ztrátovou funkci portfolia nyní označme $g(\lambda,R) = -R^T\lambda$, která je funkcí vektoru λ

a náhodného vektoru R. Mají-li náhodné výnosy R spojité rozdělení s hustotou $p(R)$, pak platí

$$\Psi(\lambda,\zeta) = P[g(\lambda,R) \leq \zeta] = \int\limits_{g(\lambda,R)\leq\zeta} p(R)dR,$$

$$P[g(\lambda,R) \geq VaR_\alpha(\lambda)] = 1 - \alpha,$$

$$CVaR_\alpha(\lambda) = \frac{1}{1-\alpha} \int\limits_{g(\lambda,R)\geq VaR_\alpha(\lambda)} g(\lambda,R)p(R)dR.$$

K praktickému využití pro naši úlohu nás bude zajímat situace, kdy se bude jednat o diskrétní rozdělení ztrát, respektive výnosů. Jak již bylo zmíněno, podrobněji se správnými výpočty CVaR$_\alpha$ zabývá text Rockafellar a Uryasev (2000), odkud jsme převzali následující rovnosti a tvrzení. Zavedeme funkci

$$F_\alpha(\lambda,\xi) = \xi + \frac{1}{1-\alpha} E\left[[g(\lambda,R) - \xi]^+\right],$$

kde $[t]^+ = max\{0,t\}$. Je možné ukázat, že platí rovnost

$$CVaR_\alpha(\lambda) = \min_\xi F_\alpha(\lambda,\xi).$$

Pro náš případ, kdy je rozdělení diskrétní, uvedeme minimalizační formuli.

Tvrzení 5. *Má-li vektor R diskrétní rozdělení v konečně mnoha bodech R_s s pravděpodobností p_s, pro $s = 1,2,\ldots,S$. Potom $CVaR_\alpha(\lambda)$ je optimální hodnotou úlohy*

$$CVaR_\alpha(\lambda) = \min_\xi \left[\xi + \frac{1}{1-\alpha} \sum_{s=1}^{S} p_s[-R_s^T\lambda - \xi]^+\right]. \tag{3.3}$$

Pokud opět uvážíme rovnoměrné rozdělení výnosů sledovaných aktiv, dostáváme vyjádření

$$CVaR_\alpha(\lambda) = \min_\xi \left[\xi + \frac{1}{1-\alpha} \cdot \frac{1}{S} \sum_{s=1}^{S} [-R_s^T\lambda - \xi]^+\right]. \tag{3.4}$$

Toto vyjádření pro nás bude klíčové a budeme ho využívat v našem diverzifikačním modelu analýzy obalu dat ve financích.

4 Diverzifikační model s aplikací podmíněné hodnoty v riziku

V kapitole druhé jsme uvedli základní myšlenky diverzifikace. Nastínili jsme, čím se diverzifikace zabývá a co budeme požadovat po užívaných mírách vstupů a výstupů. Kapitola třetí se zabývala problematikou měr rizika, konkrétně podmíněnou hodnotou v riziku ($CVaR_\alpha$). Uvedli jsme, jak se správně počítá pro případy diskrétního rozdělení. S těmito informacemi z kapitoly 2 a 3 nyní budeme pracovat. Zavedeme výsledný model, který dále užijeme pro praktickou ukázku na konkrétních datech.

Podívejme se na optimalizační úlohu obecného diverzifikačního modelu (2.1). Nyní bude naším úkolem obecné míry F_i a G_l nahradit vhodnými mírami, aby splnily požadované podmínky. Za míru rizika F_i volme CVaR_{α_i} tak, aby platilo $\alpha_i \geq 0.6$ a za G_l míru výstupů $-\mathrm{CVaR}_{\alpha_l}$ s požadavkem $\alpha_l \leq 0.4$. Pro výpočet použijeme rovnosti (3.2) a (3.4), které jsme ukázali.

Minimalizace

$$\Theta_q, \tag{4.1}$$

za podmínky

$$\begin{aligned}
-\nu_l - \frac{1}{1-\alpha_l}\cdot\frac{1}{S}\sum_{s=1}^{S}\left[-\sum_{j=1}^{n}\lambda_j R_{js} - \nu_l\right]^{+} &\geq -CVaR_{\alpha_l}(R_q), \qquad l = 1,2,\ldots,k,\\
\xi_i + \frac{1}{1-\alpha_i}\cdot\frac{1}{S}\sum_{s=1}^{S}\left[-\sum_{j=1}^{n}\lambda_j R_{js} - \xi_i\right]^{+} &\leq \Theta_q CVaR_{\alpha_i}(R_q), \qquad i = 1,2,\ldots,m,\\
\sum_{j=1}^{n}\lambda_j &= 1,\\
\lambda_j &\geq 0, \qquad j = 1,2,\ldots,n.
\end{aligned}$$

Rozhodovacími proměnnými v tomto modelu jsou Θ_q, ξ_i a λ_j pro $i = 1,\ldots,m$ a $j = 1,\ldots,n$, kde λ_j je váha j-tého aktiva a n je počet aktiv, která bereme v potaz jako investiční příležitosti. Aktuálně sledované aktivum je označeno jako q-té. Abychom zjistili, jak na tom jsou s efektivitou všechna aktiva, řešení modelu musí proběhnout n-krát. Zopakujme, že nespornou výhodou DEA modelů je možnost využití vícero vstupů. V dané úloze se toto projeví přes α_i a α_l, za které dosazujeme vhodné hodnoty z intervalu $(0,1)$ pro $i = 1,\ldots,m$ a $l = 1,\ldots,k$. Generujeme tak více vstupů, přesněji uvažujeme m různých CVaR_{α_i} a k různých CVaR_{α_l}. Vektor výnosů je reprezentován vektorem R a pro každé aktivum nese

S pozorování. Využití míry CVaR_α na různých hladinách nám zajistí, že lépe zmapujeme, jak se dané výnosy chovají, neboť různé hladiny CVaR_α zohledňují různé části rozdělení výnosů.

Jako další míra výstupu se nabízí střední hodnota výnosů. Tím pádem by nám v předchozí optimalizační úloze přibyla podmínku týkající se míry výstupů, podmínka o střední hodnotě dat

$$E\left(\sum_{j=1}^{n} \lambda_j R_j\right) \geq E(R_q). \tag{4.2}$$

Ovšem střední hodnota odpovídá míře $-\mathrm{CVaR}_\alpha$ pro $\alpha = 0$. Toto souhlasí s návrhem, že za výstupy užijeme $-\mathrm{CVaR}_\alpha$ pro $\alpha \leq 0.4$. Tím pádem není třeba model nijak měnit.

5 Aplikace modelu na data

V této části uvedeme, jak se odvozený diverzifikační model používá na reálných datech. Také ukážeme, jak by výpočet vypadal, kdybychom použili klasický DEA model. Následně porovnáme vypočtené výsledky obou optimalizačních úloh. Analýza obalu dat je velice flexibilní metoda a dala by se využít na různé druhy investic. My výpočet provedeme na podílových fondech. Podílové fondy jsou investice, kdy podílník nakupuje podílové listy a fond následně takto vložené peníze dále zhodnocuje na finančním trhu. Libovolně jsme vybrali deset zástupců v rámci českého investičního trhu, a to konkrétně

1. PROSPERITA - OPF Globální
2. Pioneer Absolute Return Control
3. J&T KOMODITNÍ FOND
4. ISČS - DLUHOPISOVÝ FOND
5. ING (L) Invest Banking & Insurance
6. ČSOB Archipel Portfolio Pro May 90
7. ČP INVEST Akciový investiční program
8. Conseq Akciový A
9. AKRO globální akciový fond
10. IKS KB privátní správa aktiv 5D - třída A.

Zvolená aktiva byla vybrána ze serveru www.kurzy.cz, odkud jsme čerpali i potřebná data. Podílové fondy jsou rozděleny do typů podle toho, jak dále investují. V našem případě jsme zvolili jeden dluhopisový fond ISČS - DLUHOPISOVÝ FOND, jeden fond fondů J&T KOMODITNÍ FOND, tři akciové ING (L) Invest Banking & Insurance, Conseq Akciový A a IKS KB privátní správa aktiv 5D - třída A. Ostatní fondy jsou vedené jako smíšené, i když ne vždy tomu jejich název odpovídá. Akciové fondy se obecně vyznačují spíše vyšším rizikem a středním až vyšším ziskem, smíšené fondy středním rizikem i středním ziskem a fondy fondů a dluhopisové spíše nižším až středním rizikem i ziskem. Náš vzorek je ale velice malý na to, abychom tato tvrzení mohli potvrdit nebo vyvrátit.

Jako data jsme použili týdenní výnosy za půl roku na trhu, tedy zhruba od poloviny září roku 2013 do poloviny března roku 2014. Uvažujeme diskrétní rozdělení výnosů R_{js} pro $s = 1,2,\ldots,S$ se shodnou pravděpodobností $\frac{1}{S}$ jako v článku Branda (2013). Pro naše výpočty platí $S = 27$. Bylo třeba některé vstupy a výstupy spočítat z dat předem a následně do modelu dosadit.

Jako zástupce klasické analýzy obalu dat jsme zvolili model BCC orientovaný na vstupy, tedy model s požadavkem variabilních výnosů z rozsahu. Užili jsme jej ve tvaru (1.9). Jako vstupy jsme zvolili CVaR_α na třech různých hladinách α, a to konkrétně $\mathrm{CVaR}_{0.75}$, $\mathrm{CVaR}_{0.9}$ a $\mathrm{CVaR}_{0.95}$. Příslušné hodnoty vstupů a výstupů jsme počítali vzorcem (3.2) uvedeným ve třetí kapitole. Jako výstup jsme zvolili průměrný týdenní výnos, tedy míru odpovídající $-\mathrm{CVaR}_0$. Využitý model vypadá následovně

Minimalizace

$$\Theta_q, \tag{5.1}$$

za podmínky

$$\begin{aligned}
-\sum_{j=1}^{n} \lambda_j CVaR_0(R_j) &\geq -CVaR_0(R_q), \\
\sum_{j=1}^{n} \lambda_j CVaR_{\alpha_i}(R_j) &\leq \Theta_q CVaR_{\alpha_i}(R_q), \qquad i = 1,2,\ldots,m, \\
\sum_{j=1}^{n} \lambda_j &= 1, \\
\lambda_j &\geq 0, \qquad j = 1,2,\ldots,n.
\end{aligned}$$

V tabulce 5.1 jsou uvedeny hodnoty, které jsme předem spočítali. Pro jednotlivá aktiva užíváme pouze zkrácené názvy obhospodařovatelů podílových fondů. Tyto hodnoty jsme spočítali na základě týdenní výkonnosti zvolených podílových fondů.

Aktivum	$\mathbf{CVaR}_{0.75}$	$\mathbf{CVaR}_{0.9}$	$\mathbf{CVaR}_{0.95}$	$\mathbf{CVaR}_{0}$ (%)
PROSPERITA	0,0124	0,0182	0,0234	0,4496
Pioneer	0,0023	0,0031	0,0032	-0.0322
J&T	0,0086	0,0133	0,0165	-0,0085
ISČS	0,0006	0,0008	0,0011	0,0137
ING	0,0200	0,0303	0,0358	0,3525
ČSOB	0,0115	0,0199	0,0272	0,1937
ČP INVEST	0,0122	0,0198	0,0261	0,3156
Conseq	0,0202	0,0379	0,0480	0,2000
AKRO	0,0202	0,0303	0,0359	0,2178
IKS KB	0,0168	0,0274	0,0334	0,3341

Tabulka 5.1: Míry rizika a výnosů užívané pro výpočty

Za model diverzifikační volíme taktéž takový, který je orientovaný na vstupy a splňuje podmínku VRS. Pro diverzifikační model jsme užili stejné míry – $\mathrm{CVaR}_{0,75}$, $\mathrm{CVaR}_{0,9}$ a $\mathrm{CVaR}_{0,95}$. Jako míru výstupu bereme střední hodnotu výnosů. Na základě modelu (2.1) naše úloha bude odpovídat zápisu (5.2).

Minimalizace

$$\Theta_q, \tag{5.2}$$

za podmínky

$$\begin{aligned}
E\left(\sum_{j=1}^{n} \lambda_j R_j\right) &\geq E(R_q), \\
CVaR_{\alpha_i}\left(\sum_{j=1}^{n} \lambda_j R_j\right) &\leq \Theta_q CVaR_{\alpha_i}(R_q), \qquad i = 1,2,\ldots,m, \\
\sum_{j=1}^{n} \lambda_j &= 1, \\
\lambda_j &\geq 0, \qquad j = 1,2,\ldots,n.
\end{aligned}$$

V kapitole třetí jsme ukázali, jak se správně počítají hodnoty CVaR_α a minimalizační formule pro diskrétní rozdělení výnosů a ve čtvrté kapitole jsme tyto poznatky již uvedli v konkrétním modelu. Vycházíme z modelu ve čtvrté kapitole (4.1) a podmínky (4.2). Nyní ukážeme model, který dále aplikujeme na naše vybraná data. Další možné uvedení diverzifikačních modelů bychom nalezli v článku Branda (2013).

Minimalizace

$$\Theta_q, \tag{5.3}$$

za podmínky

$$\begin{aligned}
E\left(\sum_{j=1}^{n} \lambda_j R_j\right) &\geq E(R_q), \\
\xi_i + \frac{1}{1-\alpha_i} \cdot \frac{1}{S} \sum_{s=1}^{S} \left[-\sum_{j=1}^{n} \lambda_j R_{js} - \xi_i \right]^+ &\leq \Theta_q CVaR_{\alpha_i}(R_q), \qquad i = 1,2,\ldots,m, \\
\sum_{j=1}^{n} \lambda_j &= 1, \\
\lambda_j &\geq 0, \qquad j = 1,2,\ldots,n.
\end{aligned}$$

Výpočty jsme provedli v programu **Wolfram Mathematica 9.0 Student Edition**. Zdrojový kód je uveden v příloze této práce. Výsledky dílčích výpočtů jsme následně dosadili do stěžejní funkce *NMinimize*, pomocí níž jsme dopočítali výsledné modelové výstupy. Tuto funkci bylo možné využít pro oba případy, jak pro případ klasického BCC modelu, tak pro model diverzifikační s předpokladem VRS.

BCC model zjišťoval celkem jedenáct neznámých, deset váhy pro jednotlivé podílové fondy a měřítko efektivity Θ_q. Pro model zohledňující diverzifikaci přibyly ještě proměnné ξ_i dané postupem výpočtu míry rizika CVaR_{α_i}.

Tabulka 5.2 ukazuje vypočítané hodnoty. Jsou zde uvedeny pouze ty, které vypovídají o efektivitě sledovaného aktiva. Výsledky obou modelů se samozřejmě liší a následně je budeme porovnávat. Dále v tabulce uvádíme skóre, tedy umístění jednotlivých aktiv po užití modelu.

	BCC model	**Skóre$_{BCC}$**	**Diverzifikační model**	**Skóre$_{D}$**
PROSPERITA	1,0000	1-2	0,9998	2
Pioneer	0,3523	7	0,2776	9
J&T	0,0695	10	0,0622	10
ISČS	1,0000	1-2	0,9999	1
ING	0,5138	5	0,5135	5
ČSOB	0,4729	6	0,4721	6
ČP INVEST	0,7205	3	0,8536	3
Conseq	0,2789	9	0,2786	8
AKRO	0,3221	8	0,3182	7
IKS KB	0,5517	4	0,5554	4

Tabulka 5.2: Výsledky BCC a diverzifikačního modelu s pořadím umístění aktiv

Z uvedené tabulky je zřejmé, že diverzifikační model z velké části naplnil naše očekávání. Dá se říct, že obecně je jistě matematicky silnější než klasické modely analýzy obalu dat. Mělo by platit, že efektivita vypočtená diverzifikačním modelem je menší nebo rovná efektivitě vypočtené klasickým DEA modelem. Z toho dále vyplývá, že jednotka, která je efektivní podle diverzifikačního modelu, bude efektivní i v modelu klasickém.

V případě našeho sedmého uvažovaného aktiva ČP INVEST Akciový investiční program a posledního IKS KB privátní správa aktiv 5D - třída A nastala výjimka, neboť po aplikaci diverzifikačního modelu jejich efektivnost ještě stoupla, nicméně umístění těchto aktiv zůstalo zachováno. Nemusí platit, že pořadí aktiv zůstane zachováno po aplikaci obou modelů. Ani v našem případě tomu tak nebylo. Fond Pioneer Absolute Return Control si polepšil o dvě příčky, naproti tomu fondy AKRO globální akciový fond a Conseq Akciový A propadly každý o jednu příčku.

Jednotka, která si svoji efektivnost téměř zachovala je PROSPERITA - OPF Globální. Pokud se na tento fond podíváme jako na samostatnou jednotku, v tabulce 5.1 vidíme, že vykazuje prakticky nejvyšší riziko, ale i výrazně vyšší zisk oproti dalším fondům. Taktéž fond ISČS - DLUHOPISOVÝ FOND udržel vysoké skóre. Podle tabulky 5.1 je zřejmé, že míry rizika tohoto fondu jsou na velice nízké úrovni, taktéž i míra výnosu.

Za povšimnutí stojí, že ani jeden z akciových fondů neskončil na vyšších příčkách skóre. Pouze fondy smíšené a fond dluhopisový se mohly blížit efektivnosti.

Závěr

Na začátku této práce jsme se zabývali klasickými modely analýzy obalu dat. Po běžných mean-risk modelech jsou dalším krokem v metodách hodnocení investic. Základem DEA modelů je optimalizační úloha. Tato metoda je velice univerzální a nabízí využití při hodnocení efektivity jednotek nejen v oblasti financí. Velikou výhodou je možnost volby vícero vstupů a výstupů podle potřeb aktuálně sledovaných jednotek. Požadavkem pozorovaných jednotek je shodná struktura vstupů a výstupů. Volíme ta data, která považujeme za relevantní, tak můžeme model přizpůsobit potřebám našeho bádání. Zajímala nás především aplikace v oblasti financí a později jsme volili jako pozorovaná aktiva výběr podílových fondů z českého investičního prostředí.

Tato metody má ovšem určité nedostatky. Posuzujeme-li investiční příležitosti, metoda neuvažuje možné závislosti mezi aktivy. Zabývat se tímto problémem a pokusit se ho vyřešit bylo předmětem této práce. Zavedli jsme diverzifikační model, tím byly tyto případné závislosti zohledněny. Model porovnává vstupy a výstupy jednotlivých aktiv se vstupy a výstupy všech možných portfolií. Dále jsme navrhli vhodné předpoklady pro míry rizika a míry výnosu tak, aby úloha, která bere v potaz diverzifikaci, byla lépe řešitelná.

Požadovali jsme, aby úloha byla konvexním programováním, což jsme zaručili koherentními mírami rizika a vhodnými mírami výnosů. V závěrečné aplikaci jsme uvedli model, který užívá jako míry rizika podmíněné hodnoty v riziku na různých hladinách a jako míru výnosu střední hodnotu. Obdobné míry jsme aplikovali i do klasického modelu DEA. Jako zástupce klasických modelů jsme vybrali model BCC. Následně jsme porovnali výsledky obou úloh na základě výstupů z programu Wolphram Mathematica 9.0 Student.

Diverzifikační model se ukázal, jako přísnější měřítko efektivnosti. Ve většině případů vycházeli hodnoty efektivity podílových fondů nižší než v případě modelu BCC. Jako efektivní se v obou případech ukázaly fondy PROSPERITA - OPF Globální a ISČS - DLUHOPISOVÝ FOND. Nicméně hodnoty, které byly spočítány, se mezi modely nelišily nijak zásadně a pořadí umístění aktiv mezi modely si bylo též velmi blízké.

Diverzifikační modely analýzy obalu dat jsou velice univerzální metodou, která umožňuje hodnotit efektivnost nejen podílových fondů, ale zajisté i mnoha jiných aktiv. Zároveň stále nabízí prostor ke zkoumání a zlepšování modelu tak, aby hodnotil investiční příležitosti co nejspolehlivěji a poskytoval tak co nejideálnější metodu sledování efektivnosti v rámci portfolia.

Literatura

Banker, R. D., Charnes, A. a Cooper, W. W. (1984). Some models for estimating techtical and scale inefficiencies in data envelopment analysis. *Management Science*, (9), 1078–1092.

Branda, M. (2013). Diversification-consistent DEA-risk test–solution techniques and an empirical comparison. *Proceedings of the 31st International Conference on Mathematical Methods in Economics*, (1), 77–82.

Branda, M. a Kopa, M. (2012). Dea-risk efficiency and stochasic dominance efficiency of stock indices. *Czech Journal of Economics and Finance*, (2), 106–124.

Charnes, A., Cooper, W. W. a Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, (2), 429–444.

Dupačová, J. a Lachout, P. (2011). *Úvod do optimalizace*. První vydání. Matfyzpress, Praha. ISBN 978-807-3781-767.

Jablonský, J. (2011). *Programy pro matematické modelování*. Oeconomics, Praha. ISBN 978-80-245-1810-7.

Lamb, J. D. a Tee, K. H. (2012). Data envelopment analysis models on investment funds. *European Journal of Operational Research*, (216), 687–696.

Lozano, S. a Gutiérrez, E. (2008). Data envelopment analysis of mutual funds based on second-order stochastic dominance. *European Journal of Operational Research*, (189), 230–244.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*, (1), 77–91.

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

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

Sůva, P. (2009). Optimalizace portfolia a míry rizika. *Referát, Matematicko-fyzikální fakulta Univerzity Karlovy.*

Seznam obrázků

Seznam tabulek

Seznam použitých zkratek

BCC = Banker, Charnes, Cooper
CCR = Charnes, Cooper, Rhodes
CRS = Constant Returns of Scale *(konstantní výnosy z rozsahu)*
CVaR = Conditional Value at Risk *(podmíněná hodnota v riziku)*
DEA = Data Envelopment Analysis *(analýza obalu dat)*
VaR = Value at Risk *(hodnota v riziku)*
VRS = Variable Returns of Scale *(variabilní výnosy z rozsahu*)

Přílohy

Zdrojový kód z programu Wolfram Mathematica 9.0 Student

Načtení týdenních výnosů vybraných podílových fondů

```
SetDirectory[NotebookDirectory[]]
R = ToExpression[Import["Výnosy podílových fondů.txt"]]/100
```

Seřazení dat a převedení na ztrátovou funkci

```
Z = Table[Sort[-R[[u]]], {u, 1, Length[R]}]
```

Výpočet indexů pro jednotlivé hladiny α

```
alfa = {0, 0.75, 0.9, 0.95}
S = Length[R[[1]]]
index = First[ Transpose[ Table[NMinimize[{a, (a - 1) < alfa[[i]]*S
&& a ≥ alfa[[i]]*S && a ∈ Integers}, {a}],
   {i, 1, Length[alfa]}]]]
```

Výpočet CVaR_α na různých hladinách α dle (3.2)

```
CVaR = Transpose[ Table[Table[ 1/(1 - alfa[[l]])*((index[[l]]/S -
alfa[[l]])* Z[[t]][[index[[l]]]] + (1/S)*Sum[Z[[t]][[j]], {j,
index[[l]] + 1, S}]), {l, 1, Length[alfa]}], {t, 1, Length[R]}]]
```

Model BCC (1.9)

```
lambdy = {la1, la2, la3, la4, la5, la6, la7, la8, la9, la10}
Table[NMinimize[{theta, Sum[lambdy[[i]]*CVaR[[2]][[i]], {i, 1,
Length[R]}] ≤ theta*CVaR[[2]][[q]] && Sum[lambdy[[i]]*CVaR[[3]][[i]],
{i, 1, Length[R]}] ≤ theta*CVaR[[3]][[q]] && Sum[lambdy[[i]]*
CVaR[[4]][[i]], {i, 1, Length[R]}] ≤ theta*CVaR[[4]][[q]] &&
-Sum[lambdy[[i]]*CVaR[[1]][[i]], {i, 1, Length[R]} ≥ -CVaR[[1]][[q]]
&& Total[lambdy] == 1 && la1 ≥ 0 && la2 ≥ 0 && la3 ≥ 0 && la4 ≥ 0
&& la5 ≥ 0 && la6 ≥ 0 && la7 ≥ 0 && la8 ≥ 0 && la9 ≥ 0 && la10 ≥
0}, Join[{theta}, lambdy]], {q, 1, Length[R]}]
```

Diverzifikační model (5.3)

```
Table[NMinimize[{theta, xi + (1/(1 - alfa[[2]]))*(1/S)*Sum[Max[0,
Sum[lambdy[[j]]*Z[[j]][[s]], {j, 1, Length[R]}] - xi], {s, 1, S}] ≤
theta*CVaR[[2]][[q]] && xa + (1/(1 - alfa[[3]]))*(1/S)*Sum[Max[0,
Sum[lambdy[[j]]*Z[[j]][[s]], {j, 1, Length[R]}] - xa], {s, 1, S}] ≤
theta*CVaR[[3]][[q]] && xo + (1/(1 - alfa[[4]]))*(1/S)*Sum[Max[0,
Sum[lambdy[[j]]*Z[[j]][[s]], {j, 1, Length[R]}] - xo], {s, 1, S}] ≤
theta*CVaR[[4]][[q]] && Mean[Sum[lambdy[[j]]*R[[j]], {j, 1,
Length[R]}]] ≥ Mean[R[[q]]] && Total[lambdy] == 1 && 0 ≤ thetaa ≤ 1
&& la1≥ 0 && la2 ≥ 0 && la3≥ 0 && la4 ≥ 0 && la5 ≥ 0 && la6 ≥ 0 &&
la7≥ 0 && la8 ≥ 0 && la9 ≥ 0 && la10≥ 0}, { {theta, 0, 1}, {xi, -1,
1}, {xa, -1, 1}, {xo, -1, 1}, {la1, 0, 1}, {la2, 0, 1}, {la3, 0, 1},
{la4, 0, 1}, {la5, 0, 1},{la6, 0, 1}, {la7, 0, 1}, {la8, 0, 1}, {la9,
0, 1}, {la10, 0, 1} }, Method → { "DifferentialEvolution",
"InitialPoints"→ { {0, -1, -1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
{0.5, 0, 0, 0, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5}, {1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1} }, "SearchPoints"→ 5,
"RandomSeed"→ 1, "ScalingFactor"→ 0.6}], {q, 1, Length[R]}]
```