

Katedra ekonometrie  
Fakulty informatiky a statistiky VŠE v Praze



a

Katedra operačného výskumu a ekonometrie  
Fakulty hospodárskej informatiky EU v Bratislave



a

Slovenská spoločnosť pre operačný výskum

**MEZINÁRODNÍ VĚDECKÝ SEMINÁŘ  
NOVÉ TRENDY V EKONOMETRII  
A OPERAČNÍM VÝZKUMU**

Katedry ekonometrie FIS VŠE v Praze

a

Katedry operačného výskumu a ekonometrie FHI EU v Bratislave

a

Slovenskej spoločnosti pre operačný výskum

ZBORNÍK

29. november – 1. december 2017

29. listopad – 1. prosinec 2017

Praha

### **Programový výbor – recenzenti:**

prof. Ing. Ivan Brezina, CSc., EU Bratislava  
doc. Ing. Mgr. Martin Dlouhý, Dr. MSc., VŠE Praha  
prof. RNDr. Ing. Petr Fiala, CSc. MBA., VŠE Praha  
prof. Ing. Josef Jablonský, CSc., VŠE Praha  
doc. Ing. Martin Lukáčik, PhD., EU Bratislava  
prof. Mgr. Juraj Pekár, PhD., EU Bratislava

### **Organizačný výbor:**

Ing. Jan Zouhar, PhD.  
Ing. Brian König, PhD.

**Náklad: 50 ks**  
**AH 10,07**

**Kontakt:** [brian.konig@euba.sk](mailto:brian.konig@euba.sk)

**Editori:** prof. Ing. Josef Jablonský, CSc. , Ing. Brian König, PhD.

**Web editor:** doc. Ing. Martin Lukáčik, PhD.

Za odbornú stránku a jazykovú úpravu zodpovedajú autori.

Fakulta informatiky a statistiky VŠE v Prahe,  
nám. W. Churchilla 4, 130 67 Praha 3  
a Fakulta hospodárskej informatiky EU v Bratislave,  
Dolnozemska cesta 1, 852 35 Bratislava

© Ekonomická univerzita v Bratislave,  
Vydavateľstvo EKONÓM, Bratislava 2017

ISBN: 978-80-225-4455-9

# Obsah

<i>Ivan Brezina, Ladislav Varga:</i> <a href="#">Využitie programu PrimaveraP6 pri plánovaní rozoberania likvidovaných vozidiel</a>	5
<i>Marek Dvořák:</i> <a href="#">Senzitivita hyperparametrů CNN modelu časové řady</a>	12
<i>Dana Figurová:</i> <a href="#">Modely a algoritmy na riešenie úlohy kuriéra</a>	17
<i>Pavel Gežík:</i> <a href="#">Využitie princípov vzájomného zápočtu záväzkov vo využívaní energie z obnoviteľných zdrojov</a>	25
<i>Vladimír Holý:</i> <a href="#">Kvadratické estimátory kvadratické variace</a>	32
<i>Eduard Hozlár:</i> <a href="#">Modelovanie optimálnej štruktúry financovania podniku pri neúplnej informácii</a>	40
<i>Michaela Chocholatá:</i> <a href="#">Pretrvávajúce regionálne disparity na Slovensku</a>	47
<i>Jaroslav Janáček:</i> <a href="#">Strategie „nejbližší bere“ v umisťovacích úlohách</a>	55
<i>Miroslava Jánošová:</i> <a href="#">Analýza zložiek ukazovateľa medzinárodnej investičnej pozície a ich vplyv na hrubý domáci produkt</a>	60
<i>Rudolf Kucharčík, Zuzana Čičková:</i> <a href="#">Potenciálne koalície po parlamentných voľbách v rokoch 2012 a 2016 v slovenskej republike</a>	67
<i>Martin Lukáčik:</i> <a href="#">Využitie metódy maximálnej vierohodnosti v programe EViews</a>	75
<i>Adriana Lukáčiková:</i> <a href="#">Odhad ceny bytov využitím ekonometrického modelu</a>	81
<i>Simona Macková:</i> <a href="#">Prostorová ekonometrie v makroekonomické analýze</a>	85

<i>Renata Majovská, Petr Fiala:</i> <a href="#">Cenové modely revenue managementu</a>	93
<i>Veronika Miřková:</i> <a href="#">Produkčné údaje v matici spoločenského účtovníctva</a>	99
<i>Vladimír Mlynarovič:</i> <a href="#">Algebra viacfaktorovej úlohy výberu portfólia</a>	106
<i>Juraj Pekár:</i> <a href="#">Technológia spracovania automobilového odpadu a životný cyklus automobilov</a>	112
<i>Michal Pieter:</i> <a href="#">Analýza obalu dát ako nástroj viackriteriálneho hodnotenia variantov</a>	118
<i>Tatiana Polonyankina:</i> <a href="#">Mosaikové grafy jako nástroj pro vizualizaci dat a tvorbu modelu: případová studie pro španělský pracovní trh</a>	126
<i>Allan Jose Sequeira Lopez:</i> <a href="#">Priestorová diferenciácia a teória hier</a>	133
<i>Ondřej Sokol:</i> <a href="#">Výběrový rozptyl nad intervalovými daty: srovnání optimalizačních algoritmů v MATLABu</a>	138
<i>Karol Szomolányi:</i> <a href="#">Pavučinový model so zásobami</a>	146
<i>Michal Švarc:</i> <a href="#">Předzpracování dat v oblasti retailu</a>	151
<i>Petra Tomanová:</i> <a href="#">Porovnání odhadů stochastické volatility: analýza S&amp;P 500</a>	158
<i>Petra Zýková:</i> <a href="#">Scoring Rules - DEA/AR model s penalizací</a>	164

# VYUŽITIE PROGRAMU PRIMAVERAP6 PRI PLÁNOVANÍ ROZOBERANIA LIKVIDOVANÝCH VOZIDIEL

## USE PROGRAM PRIMAVERAP6 IN PLANNING DISMANTLING DISPOSED VEHICLES

*Ivan Brezina, Ladislav Varga*

### **Abstrakt**

Projektový manažment je v súčasnosti veľmi často zaužívaný pojem obzvlášť v stredných a veľkých firmách, v ktorých sú využívané princípy a metódy projektového riadenia k úspešnej realizácii projektov. Nadväznosť jednotlivých procesov pri likvidácii starého vozidla môže byť základom pre optimalizáciu jednotlivých na seba nadväzujúcich činností a teda aj na využitie princípov projektového manažmentu. V príspevku je opísaný softvérový produkt PrimaveraP6 od spoločnosti Oracle. Tento výkonný softvér je určený stredným a väčším firmám, ktoré kladú vysoké nároky na projektové riadenie.

***Kľúčové slová:** Primavera P6, likvidácia starých vozidiel, projektový manažment*

### **Abstract**

Project management is currently often used term especially in medium and large firms, where project management principles and methods are used to successfully implement projects. Traceability of individual processes in the disposal of old vehicles can be the basis for optimizing each of interlinked activities and thus to use of the principles of project management. The paper describes Oracle's PrimaveraP6 software product. This is a powerful software is intended for medium-sized and larger businesses that place high demands on project management.

***Keywords:** Primavera P6, disposal of old vehicles, project management*

## **1 ÚVOD**

V súčasnosti existuje viac ako 300 softvérových riešení pre plánovanie a riadenie projektov. Z tohto faktu plynie, že už len rozhodnutie, pre ktorý softvér sa firma rozhodne, sa stáva vlastne projektom. V posledných rokoch sa na trhu rapídne zvyšuje podiel webovo orientovaného softvéru pre plánovanie a riadenie projektov, t.j. softvéru, ktorý umožňuje prístup k dátam projektu prostredníctvom internetového prehliadača.

Samozrejmosťou pri využívaní softvéru pre podporu, plánovanie a riadenie projektov je možnosť súčasnej práce viacerých používateľov na viacerých projektoch. Väčšina softvérov, ktoré ponúkajú túto možnosť, vyžaduje inštaláciu serverovej komponenty. Každý lepší softvér pre podporu, plánovanie a riadenie projektov by mal podporovať tvorbu Ganttových diagramov, sieťových diagramov, alokáciu zdrojov, analýzu kritickej cesty, sledovanie postupu v projekte, vytváranie čiastkových správ o stave projektu a tiež poskytnúť možnosť integrácie s databázovým softvérom minimálne na podnikovej úrovni. Softvérové licencie k týmto produktom sa spravidla predávajú každému záujemcovi „na mieru“, t.j. podľa počtu používateľov, veľkosti a zamerania firmy, prípadne aj na časovo obmedzené obdobie.

Najväčší podiel na trhu so softvérom pre plánovanie a riadenie projektov metódou CPM má spoločnosť Microsoft, s produktom MS Project a spoločnosť Oracle, s produktom PrimaveraP6. Najvhodnejším softvérom pre riadenie projektov metódou CCPM je softvér Concerto Projects, ktorý bol vyvinutý vývojármi Concerto Integrated Software Solutions v roku 2004.

Ekologické spracovanie vozidiel s ukončenou životnosťou sa v posledných rokoch stali aktuálnym problémom. Asi štyri pätiny, hlavne kovových materiálov, je možné recyklovať, znovu použiť alebo regenerovať, s jednou pätinou je problém. Úprava zvyšného odpadu na recykláciu nie je ekonomická, lacnejšie je uložiť ho na skládkach. Plasty vo zvýšenej miere nahrádzajú kovy, pretože znižujú hmotnosť automobilu, a tým jeho spotrebu paliva. Na rozdiel od kovov sa však plasty recyklujú oveľa ťažšie.

Medzi hlavné prístupy, ktoré sa v súčasnosti využívajú pri spracovaní starých vozidiel, patria:

- Šrédrovanie (americký prístup), ktorý predstavuje najstarší spôsob spracovania starých vozidiel, je ich drvenie, pričom sú staré automobily komplexne zošrotované vrátane motorov, pneumatík, čalúnených sedadiel, osí automobilov a pod., následne sú z tejto hmoty separované kovové časti od nekovových.
- Selektívna demontáž (francúzsky prístup), ktorú predstavuje predbežná demontáž a následné šrédrovanie.
- Úplná demontáž (nemecký prístup), ktorej cieľom je 100 % recyklácia starých vozidiel.

Najlepšie prepracovaný systém spracovania starých vozidiel v Európe má firma Renault. Po celom území štátu je vytvorená sieť zberných organizácií, kde sa staré, ojazdené, havarované a vyradené vozidlá zhromažďujú a následne v demontážnych strediskách sa recyklovateľné súčiastky recyklujú (napr. nárazníky na znovuzískanie plastov) a celá demontáž sa robí podľa demontážnej knihy či návodu, ktorý je rôzny podľa značky a typu auta. Použiteľné súčiastky idú do predajne náhradných dielov. Očistená karoséria vozidla ide na šrédrovanie. Na magnetickom separátore sa oddelí podiel železa, medi a hliníka, zvyšok, ktorý tvoria plasty, textil (poťahy sedadiel), guma a sklo sa separuje podľa zrnitosti.

Práve nadväznosť jednotlivých procesov môže byť základom pre optimalizáciu jednotlivých na seba nadväzujúcich činností.

## **2 PROJEKTOVÝ MANAŽMENT**

Pri súčasnej tendencii skracovania životného cyklu projektu a pri snahe o koordináciu zdrojov a aktivít by sa mali nástroje a techniky projektového riadenia uplatňovať vo väčšom rozsahu. Hromadné preniknutie projektového riadenia do riadiacej praxe vyžaduje pochopenie a záujem všetkých účastníkov projektového riadenia. Nástroje projektového riadenia poskytujú flexibilitu pre plánovanie, riadenie a tiež sledovanie projektov. Dávajú možnosť rýchle a efektívne reagovať na nevyhnutné zmeny v projektoch. Sofistikované postupy sa používajú pre časovo limitované a zdrojovo obmedzené rozvrhovanie a tiež pre realizáciu viacerých projektov súčasne. Tento prístup založený na využívaní zdrojov pre dosiahnutie cieľov v požadovanom čase a v požadovaných nákladoch by mal byť užitočný pre všetkých manažérov.

Postupne sa prejavuje trend využívania projektov pre riadenie firiem, takzvané projektové riadenie. Projektové riadenie ako manažérska stratégia projektovo orientovaných firiem

poskytuje možnosť bohatej budúcnosti. Projektovému riadeniu sa v širokej miere venuje Hrazdilová Bočková v publikácii Projektové řízení (Hrazdilová Bočková, 2016).

Podľa Fialu (Fiala, 2004) je projektové riadenie spôsob riadenia prostredníctvom projektov. Projektové riadenie predstavuje veľmi účinný nástroj riadenia zmien a komplexnú koncepciu efektívneho dosahovania strategických cieľov. Projektové riadenie umožňuje manažérom dosiahnuť kvalitný výstup s minimálnymi nárokmi na čas a ostatné zdroje. Projektové riadenie zahŕňa riadenie jednotlivých projektov, vytvorenie organizačnej štruktúry, koordináciu projektov z hľadiska časových termínov a disponibilných zdrojov.

Nástroje projektového riadenia poskytujú flexibilitu pre plánovanie, riadenie a sledovanie projektov. Dávajú možnosť rýchle a efektívne reagovať na nevyhnutné zmeny projektov. Pre časovo limitované a zdrojovo obmedzené rozhodovanie viac rôznych typov projektov súčasne sa v dnešnej dobe používajú sofistikované postupy a nástroje. Projektové riadenie je štýl riadenia, ktorý by sa mal v súčasných dynamických podmienkach presadzovať stále vo väčšej miere.

V súčasnosti existuje niekoľko pohľadov na projektové riadenie. K najrozšírenejším patria manažérsky pohľad a systémovo analytický pohľad. Manažérsky pohľad vychádza z klasického manažmentu a kladie dôraz na praktickú realizáciu projektu. Rieši problémy týkajúce sa zodpovednosti pracovníkov za jednotlivé úlohy, výberu manažérov rôznych úrovní riadenia, ako aj ostatných ľudských zdrojov, motivácie pracovníkov alebo spôsobu odmeňovania. Systémovo analytický pohľad vychádza z matematickej podstaty tvorby osnovy a zaoberá sa vytváraním matematických modelov pre riadenie projektov.

### 3 METÓDY PROJEKTOVÉHO MANAŽMENTU

Štruktúry sieťových grafov ako reprezentantov realizovaných projektov môžu byť buď deterministické, alebo stochastické. V prípade zovšeobecnených sieťových grafov sa používa pravdepodobnostné ohodnotenie, ktoré udáva podmienené pravdepodobnosti pre realizáciu jednotlivých činností.

Ohodnotenie sieťového grafu môže byť tiež deterministické, alebo stochastické. V prípade deterministického ohodnotenia grafu je hodnota parametra presne určená. Keď hovoríme o stochastickom ohodnotení grafu, tak hodnotu parametra považujeme za náhodnú veličinu so známym, poprípade s neznámym pravdepodobnostným rozdelením.

V tabuľke 1 sa uvádzajú štruktúry sieťových grafov:<sup>1</sup>

Sieťový graf	Štruktúra grafu	Interpretácia ohodnotenia	Druh ohodnotenia	Metóda
Hranovo orientovaný	Deterministická	Čas	Deterministické	CPM
			Stochastické	PERT
		Náklady	Deterministické	CPM/COST
		Zdroje	Deterministické	Sumarizácia, rozvrhovanie a vyrovnanie
	Stochastická	Čas	Deterministické	GERT
		Náklady	Deterministické	
Uzlovo definovaný	Deterministická	Čas	Deterministické	MPM
			Stochastické	Počítačové systémy
		Náklady	Deterministické	
		Zdroje	Deterministické	

Tabuľka 1: Štruktúry sieťových grafov – vlastná práca v programe MS Excel, zdroj: Brezina, I. – Čičková, Z. – Gežík, P. Sieťová analýza. 2012.

<sup>1</sup> BREZINA, I. – ČIČKOVÁ, Z. – GEŽÍK, P. 2012. Sieťová analýza. Vydavateľstvo ekonóm 2012. ISBN 978-80-225-3503-8.

Podľa zvolenej interpretácie sieťového grafu, štruktúry grafu, interpretácie ohodnotení činností a tiež ich druhov získavame množstvo možných metód vhodných k danej analýze:

- Metóda kritickej cesty – CPM (Critical Path Method).
- Metóda hodnotenia a kontroly projektu – PERT (Project Evaluation and Review Technique). Metóda PERT používa stochastické ohodnotenie sieťového grafu, založená je tiež na určenie kritickej cesty projektu.
- Metóda MPM (Metra Potencial Method). Metóda MPM pracuje s uzlovo ohodnotenými sieťovými grafmi, u ktorých sa predpokladá deterministické ohodnotenie a tiež deterministická štruktúra sieťového grafu.
- Metóda kritického reťazca – CCPM (Critical Chain Project Management). Metóda CCPM je založená na priamej aplikácii teórie obmedzenia (TOC).
- Metóda grafického ohodnotenia a kontroly projektu – GERT (Graphical Evaluation and Review Technique). Metóda GERT pracuje so zovšeobecneným sieťovým grafom.
- Metóda PDM (Precedence Diagramming Method). V prípade metódy PDM sa jedná o jednoduchšiu variantu metódy MPM.

## 4 PROGRAM PRIMAVERA P6

PrimaveraP6 je výkonný softvér od spoločnosti Oracle, ktorý je určený stredným a väčším firmám, ktoré kladú vysoké nároky na projektové riadenie. Pomáha im spracovať kompletne projektové portfólio v priebehu ich životnosti. Softvér PrimaveraP6 je určený na spracovanie rozsiahlych projektov, tiež umožňuje firmám zlepšovať rozhodovanie sa pri riadení portfólia projektu, vyhodnocovať riziká aj prínosy spojené s jednotlivými projektmi a rozhodovať sa, či je k dispozícii dostatok zdrojov pre prevedenie práce. PrimaveraP6 poskytuje jedinečné riešenie pre riadenie projektov akejkoľvek veľkosti, prispôsobí sa rôznym úrovňam zložitosti v rámci projektu a je inteligentne upravený tak, aby vyhovoval potrebám rôznych rolí, funkcií či úrovni vo firme aj celému projektovému tímu. Ponúka sofistikované a flexibilné nástroje pre organizovanie, riadenie, filtrovanie aktivít, zdrojov a tiež celých projektov.

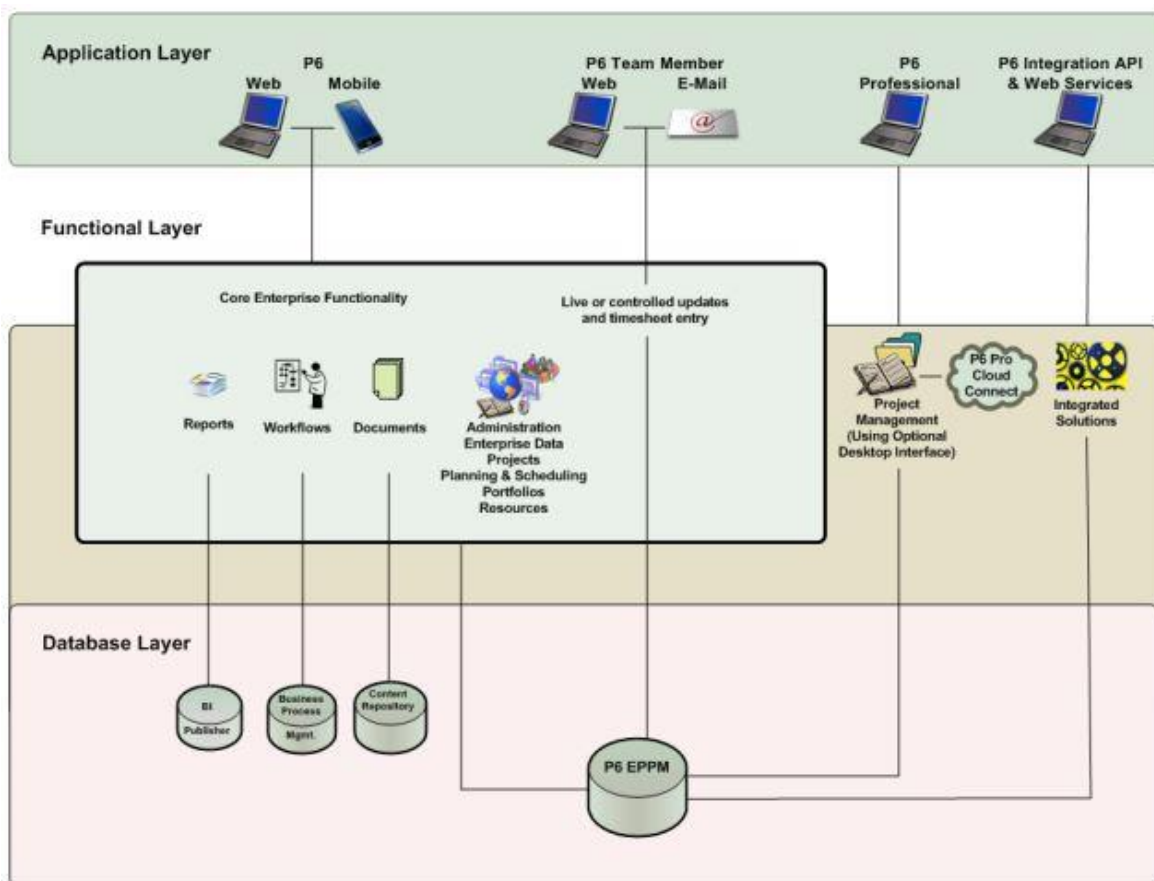
PrimaveraP6 spolupracuje s ďalšími voliteľnými softvérovými riešeniami ako sú Primavera Contact Management, Oracle Primavera Prime, Oracle Primavera Risk Analysis, Primavera Unifier, Oracle PPM a Primavera Earned Value Management. V závislosti na nasadení softvéru P6 v konkrétnej firme, sa architektúra všeobecne skladá z aplikačnej vrstvy, funkčnej vrstvy a z databázovej vrstvy. Architektúra P6 je znázornená na obrázku 1.

Aplikačná vrstva softvéru PrimaveraP6

- P6 Web – väčšina používateľov spolieha takmer výlučne len na P6 webovú aplikáciu, ktorá je spustiteľná v štandardnom internetovom prehliadači. Jednoducho povedané P6 Web je primárne rozhranie pre správu a riadenie projektov.
- P6 Mobile – umožňuje členom projektového tímu získať informácie o stavoch úloh, prezeranie harmonogramu projektu a tiež umožňuje prácu „na cestách“.
- P6 Team Member Web – toto webové rozhranie umožňuje členom projektového tímu získať informácie o stave úloh a tiež časových záznamoch v časovom harmonograme pomocou prehliadača.



- P6 Team Member E-mail – táto aplikácia umožňuje členom projektového tímu poskytnúť informácie o stave úloh pomocou ľubovoľného emailového konta.
- P6 Professional – je hlavné rozhranie pre všetky projektové funkcie a správu projektu. Aplikácia P6 Professional a všetky jeho funkcie vrátane P6 Visualizer sú spustiteľné na operačných systémoch Microsoft Windows. P6 Professional je k dispozícii aj pre používateľov, ktorí potrebujú pracovať na projektoch v offline režime, alebo pre používateľov pracujúcich v prostredí Cloud.
- P6 Integration API – je programové rozhranie aplikácie založené na programovacom jazyku Java, umožňujúce prepojenie s ostatnými komponentmi a systémami.
- P6 Web Services – je integrácia technológie, ktorá rozširuje funkčnosť na otvorených štandardoch ako sú napríklad SOAP, XML a WSDL. P6 Web Services umožňuje vývojárom využiť štandardné rozhranie pre vytvorenie integrovaného softvérového riešenia, ktoré obsahuje širokú škálu podnikových aplikácií spustiteľných na rôznych hardvérových a operačných systémoch.



Obrázok 1: Architektúra softvéru PrimaveraP6 – zdroj: P6 EPPM System Architecture Data Sheet, Release 8.4

#### Funkčná vrstva softvéru PrimaveraP6

- Administration and Enterprise Data – táto funkcionálna umožňuje spravovať užívateľské kontá, preferencie, pohľady, aplikačné nastavenia a podnikové dáta.

- Project, Portfolio, and Resource Management – poskytuje širokú škálu funkcií navrhnutých pre optimalizáciu všetkých fáz projektového manažmentu, riadeniu zdrojov a riadeniu portfólia. Táto funkcionalita zahŕňa plnú podporu pre aktivity, štruktúry, plánovanie a rozvrhovanie, riadenie nákladov, správu zdrojov, riadenie jednotlivých tímov, analýzu portfólia, plánovacie kapacity, návratnosti investícií a funkcie určené na meranie výkonnosti na všetkých úrovniach a v každom okamihu.
- Documents – zahŕňa podporu pre správu dokumentov s možnosťou voliteľného úložiska. Tiež zahŕňa voliteľné funkcie pre spoluprácu zúčastneným stranám na jednotlivých dokumentoch s možnosťou vykonania revízie dokumentov.
- Workflows – poskytuje BPMN (Business Process Modeling Notation) kompatibilné grafické notácie, ktoré zobrazujú pracovné postupy a kroky od začatia projektu.
- Reports – generovanie a zobrazenie zostáv pomocou štandardného internetového prehliadača. Reporty môžu byť generované a smerované prostredníctvom emailu, dajú sa uložiť do zdieľaného alebo lokálneho súboru, alebo priamo vytlačiť.
- Planning and Scheduling – ponúka veľkú sadu funkcií predovšetkým pre projektantov a plánovačov zahŕňajúcu reflektovanie projektov, projektových plánov a tiež správu dizajnu plánovaného projektu. Funkcia P6 Visualizer umožňuje vytvárať a prispôbovať vizuálne zobrazenie komplexných informácií o pláne projektu vrátane Ganttových diagramov. Niektoré funkcie sú k dispozícii iba po pripojení k EPPM databáz.
- Integrated Solutions – umožňuje vytváranie, alebo rozmiestnenie iných systémov pomocou Integration API a Web Services.

Kompletná dokumentácia a online knižnice pre softvér PrimaveraP6 sú k dispozícii na technológii Oracle Network (OTN).

## 5 ZÁVER

Programový softvér PrimaveraP6 sa javí ako efektívny nástroj na riadenie projektov vo všeobecnosti. Pretože spracovanie starých vozidiel možno charakterizovať ako súbor na seba nadväzujúcich činností, použitie tohto softvérového nástroja môže viesť k optimalizácii usporiadania jednotlivých činností tak, aby rozmontovanie a následná recyklácia starých ojazdených vozidiel bola efektívna.

### Použitá literatúra

(podľa normy STN ISO 690 / ČSN ISO 690)

1. BREZINA, I., ČIČKOVÁ, Z., GEŽÍK, P. 2012. Siet'ová analýza. Vydavateľstvo ekonóm 2012. ISBN 978-80-225-3503-8.
2. FIALA, P. 2004. Projektové řízení: modely, metody, analýzy. První vydání Praha: Professional Publishing. ISBN 80-86419-24-X.
3. GOLDRATT, E. M. 1997. Critical Chain. Great Barrington: North Rivers Press. ISBN 08-842-7153-6.
4. HRAZDILOVÁ BOČKOVÁ, K. 2016. Projektové řízení učebnice. Vydal: Martin Koláček – E-knihy jedou 2016. ISBN 978-80-7512-431-9.
5. Project Management Software. *Concerto Projects by Concerto Support Services*. <http://www.capterra.com/project-management-software/spotlight/72494/Concerto%20Projects/Concerto%20Support%20Services>.

6. Oracle. *PrimaveraP6 Enterprise Project Portfolio Management*.  
<https://www.oracle.com/applications/primavera/index.html>.
7. <https://www.sme.sk/c/2271200/auto-ako-zaujimavy-odpad.html#ixzz4xpwXQOZo>
8. <https://www.sme.sk/c/2271200/auto-ako-zaujimavy-odpad.html>

### **Kontaktné údaje**

prof. Ing. Ivan Brezina, CSc.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 821

email: [brezina@euba.sk](mailto:brezina@euba.sk)

Ing. Ladislav Varga

Úrad podpredsedu vlády SR pre investície a informatizáciu

Štefánikova 15, 811 05 Bratislava

Tel: (+421) 949 767 007

Email: [vargaladislav1@gmail.com](mailto:vargaladislav1@gmail.com)

# SENZITIVITA HYPERPARAMETRŮ CNN MODELU ČASOVÉ ŘADY

## HYPERPARAMETER SENSITIVITY OF CNN TIMESERIES MODEL

*Marek Dvořák*<sup>1</sup>

### Abstrakt

V tomto článku se zabýváme vlivem hyperparametrů, zejména velikostí konvolučních kernelů a velikosti plně propojené závěrečné skryté vrstvy, na metriky předpovědi časových řad pomocí umělých konvolučních neuronových sítí. Je ukázáno, že některé kombinace hyperparametrů poskytují lepší výsledky.

**Klíčová slova:** neuronová síť, hyperparametry, časová řada, předpověď

### Abstract

In this paper, we investigated the influence of hyperparameters, namely the size of convolutional kernels and size of fully connected final layer, on the metrics of timeseries prediction using convolutional artificial neural networks. It was shown that certain combinations of hyperparameters yield slightly better results than other and these were discussed.

**Keywords:** neural net, hyperparameters, time series, prediction

## 1 INTRODUCTION

Neural networks are a powerful tool in the toolset of machine learning. The foundations of neural networks were placed in the middle of last century in an attempt to simulate the inner workings of biological brain. But the widespread use of neural networks had to wait until a sufficient computer power was available for practical use.

Neural networks are known to be universal function approximators (Equation 1), that is given the input  $X$ , they are able to associate an output  $Y$  at arbitrary precision.

$$f : X \rightarrow Y \quad (1)$$

Neural networks consists of many neurons. These neurons themselves are simple functions too. Typical neuron can be described using equation 2

$$o_i = f_{out} \left( \sum_{j \in J} w_{ij} o_j + b_i \right), \quad (2)$$

where  $o_i$  is the output of the current neuron  $i$ ,  $o_j$  are outputs from the set  $J$  of neurons feeding into neuron  $i$ ,  $w_{ij}$  are weights assigned to inputs from neurons in the set  $J$  to the neuron  $i$ ,  $b_i$  is a bias variable of a neuron  $i$  and  $f_{out}$  is a nonlinear monotonous activation function.

---

<sup>1</sup>The research project was supported by Grant No. IGA F4/58/2017 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

This activation function is usually a sigmoid:

$$f_{out} = \frac{1}{1 + e^{-x}}, \quad (3)$$

tanh:

$$f_{out} = \tanh x, \quad (4)$$

or more recently, ReLU [2]:

$$f_{out} = \max(0, x). \quad (5)$$

These neurons are then connected together, output on one feeding into the input of another, resulting in a network graph, which has several inputs and (optionally) several outputs. By carefully optimizing weights  $w_{ij}$ , we can then approximate function 1.

## 1.1 Convolutional neural network

Convolutional network architecture is a type of regularization that uses spatial information of input data to help more precise output estimation [3][4]. The input  $N$ -dimensional data is convoluted using several convolution filters. These filters are trained to recognize certain patterns in the input. For each kernel applied on the input, a convoluted layer it output as a channel. Optional intermediate pooling layer can be present to help reduce number of variables in subsequent layers. Several such convolutional layers with convolutional can be present, each trying to discern higher level features. The whole architecture is usually ended with simple fully connected layer of neurons to help to compile all kernel information into the output.

These convolutional kernels are not chosen arbitrarily at the beginning, but are an essential part of training process. The training algorithm tries to estimate kernels which help to minimize the learning error and which provide the most precise estimations.

While the contemporary focus of convolutional neural network research focuses on image recognition, that is 2-dimensional input, they can be applied generally on  $N$ -dimensional data such as 1-dimensional time series, which is the focus of this paper.

## 2 MODEL ARCHITECTURE

We have obtained five years of daily data of three stock indices from Nasdaq - TSLA (Tesla), MCD (McDonalds) and NKE (Nike),  $n = 1257$  data points for each index in total. Preprocessing involved simple logarithmic differences.

$$\Delta index_{i,k,t} = \ln index_{i,k,t} - \ln index_{i,k,t-1}, \quad (6)$$

where  $i \in \{TSLA, MCD, NKE\}$ ,  $k \in \{OPEN, HIGH, LOW, CLOSE, VOLUME\}$  of each stock index (in a single day) and  $t$  being time index.

We have a priori chosen several base hyperparameter values, such as the number of convolutional layers being 2, size of the first layer kernels being  $k_1 = 8$ , size of the second layer kernels  $k_2 = 16$ , first layer kernel count  $m_1 = 9$ , second layer kernel count  $m_2 = 16$  and number of neurons in the last fully connected layer being 256.

Coefficients :					
Estimate	Std. Error	t value	Pr(> t )		
(Intercept)	-8.5250000	0.5152474	-16.545	6.47e-12	***
l1	0.1000000	0.0547066	1.828	0.0852	.
l2	0.0222917	0.0136766	1.630	0.1215	
flat	0.0008464	0.0005485	1.543	0.1412	

Figure 1: Influence of hyperparameters on log MSE

The observations were constructed as follows. The size of lookbehind window was calculated as  $w = k_1 + k_2 - 1$ . We could then construct  $n - 1 - w$  observations  $j$ . Each observation had independent variables  $x_{i,k,j} \in \{\Delta index_{i,k,j-w} \dots \Delta index_{i,k,j-1}\}$  for each  $i$  and a dependant variable  $y_{1,j} \in \{\Delta index_{i,CLOSE,j}\}$ .

Each price lookbehind window was then convoluted using  $m_1$  convolution kernels of size  $k_1$ . Observations regarding volume were then convoluted using its own set of  $k_1$  convolutional kernels of size  $k_1$ .

Second layer was convolution using  $m_2$  convolution kernels of size  $k_2$  on price and volume observations separately.

The resulting  $m_2$  kernels were then flattened and fed to a fully connected neural network layer with 256 neurons which then fed to a single output neuron. All neurons used the sigmoid (Equation 3) activation function. The mean square error was then minimized using Adam algorithm [1].

### 3 HYPERPARAMETER TUNING

The goal of this paper was to determine the influence of hyperparameter tuning on the model metric. From the base hyperparameters mentioned in section 2, we constructed additional models using different values. These included changing the kernel size of first convolutional layer to 6 and 10, changing the kernel size of second convolutional layer to 8 and 24 and changing the number of neurons in the last fully connected layer to 128 and 512. All with respect to base values.

The metric used was the logarithm of mean squared error.

This resulted in seven hyperparameter configurations. Three estimations were run on each configuration, 21 estimations in total.

A simple linear regression was then used to determine the influence of hyperparameter on the chosen metric.

### 4 RESULTS

Table 1 describes all run experiments using different hyperparameters.

Figure 1 shows the influence of different hyperparameters on the final logarithmic mean squared error. Lower MSE means better model fit. We can see from positive coefficients that in out particular case of neural network, the best results yields smaller network architectures.

<i>l1</i>	<i>l2</i>	<i>flat</i>	log MSE
8	16	256	-7.43
8	16	256	-6.85
8	16	256	-7.14
6	16	256	-7.17
6	16	256	-7.57
6	16	256	-7.08
10	16	256	-6.77
10	16	256	-6.74
10	16	256	-7.11
8	8	256	-7.45
8	8	256	-7.11
8	8	256	-7.75
8	24	256	-7.07
8	24	256	-7.08
8	24	256	-7.09
8	16	128	-7.01
8	16	128	-7.36
8	16	128	-7.32
8	16	512	-6.8
8	16	512	-7.44
8	16	512	-6.52

Table 1: Summary of all run experiments. *l1* is the kernel size of first convolutional layer, *l2* is the kernel size of second convolutional layer and *flat* means the number of neurons in the last fully connected hidden layer.

## 5 KERNELS

We can also take interest in the estimated kernels with the smallest (winning) architecture. These kernels are shown in figure 2.

We see that the algorithm learned several low level patterns that can appear in a time series and are capable of prediction. These 9 kernels lay at the foundation of each time series and can be thought of as building blocks.

## 6 CONCLUSION

The neural network approach for time series study is a powerful one, especially if coupled with convolutional layer for estimating time series features and behavior preceding our objective function. We have shown how it is possible to estimate convolution kernels from time series as well as how such estimated kernels look like in one such convolution layer.

We focused on hyperparameter fine tuning such as number of neurons, kernels or layers, and learned that certain combinations of hyperparameters yield better results than others.

More accurate micro and macro economic restriction could also be employed as well as wider

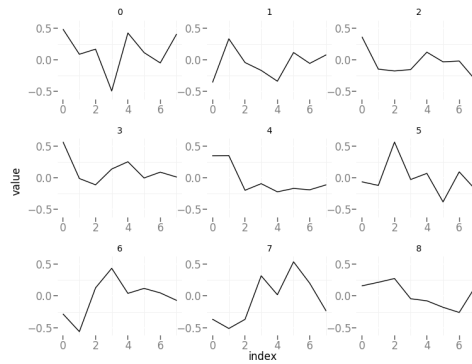


Figure 2: Estimated 9 convolution kernels (functions) of the first layer. The 'value' axis represents the kernel function value at the input position represented by the 'index' axis.

range of input data to refine prediction accuracy.

## References

- [1] KINGMA, D. a BA, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [2] LECUN, Y. A., BOTTOU, L., ORR, G. B. a MÜLLER, K.-R. (2012). Efficient backprop. In *Neural networks: Tricks of the trade*, pages 9–48. Springer.
- [3] POULTNEY, C., CHOPRA, S., CUN, Y. L. A KOL. (2007). Efficient learning of sparse representations with an energy-based model. In *Advances in neural information processing systems*, pages 1137–1144.
- [4] ZHANG, W., ITOH, K., TANIDA, J. a ICHIOKA, Y. (1990). Parallel distributed processing model with local space-invariant interconnections and its optical architecture. *Applied optics*, **29**(32), 4790–4797.

### Contact:

Ing. Marek Dvořák  
 University of Economics, Prague  
 Department of Econometrics  
 W. Churchil Sq. 4  
 130 67 Prague 3  
 Czech Republic  
 email: marek.dvorak@vse.cz



# MODELY A ALGORITMY NA RIEŠENIE ÚLOHY KURIÉRA

## MODELS AND ALGORITHMS FOR COURIER DELIVERY PROBLEM

*Dana Figurová*

### Abstrakt

Predmetom tohto príspevku je bližšie popísať úlohu kuriéra, ktorá je jednou z modifikácií známeho problému obchodného cestujúceho. Venujeme sa vymedzeniu podstaty úlohy rozvozu a zvozu, v zahraničí známou ako „*Pickup and Delivery problem*“, pod ktorú spadá aj úloha kuriéra a zároveň budeme prezentovať prístupy k jej riešeniu pomocou rôznych algoritmov a optimalizačných techník. Opísané sú rôzne typy modelov úloh kuriéra, ktoré majú široké praktické využitie.

**Príučové slová:** *úloha kuriéra, heuristické algoritmy, modifikácia obchodného cestujúceho*

### Abstract

The purpose of this study is to further describe the role of a courier, which represents one of modification of the algorithm for solving the travelling salesman problem. We are dealing with the problem, known abroad as the "Pickup and Delivery problem", which also includes courier delivery problem and at the same time we will present approaches to its solution using various algorithms and optimization techniques. Furthermore, it outlines different types of models related to the tasks of the courier, which have a world-wide practical use.

**Keywords:** *courier delivery problem, heuristic algorithms, a modification of the algorithm for solving the travelling salesman problem*

## 1 ÚVOD

Pre kuriérov je vzhľadom na veľkosť prepravovaných zásielok extrémne dôležité správne, inak povedané optimálne naplánovanie trasy, čím dokážu výrazne ušetriť časť nákladov. Dopravné okružné úlohy, ktoré sa využívajú na naplánovanie cesty takýchto firiem, používajú štandardné metódy a modely s tzv. statickým, alebo dynamickým prístupom. Statický prístup predpokladá znalosť všetkých požiadaviek od zákazníkov s predstihom, resp. značne pred tým, ako bola zahájená cesta kuriérom. V reálnych situáciách sa však stretávame skôr s dynamickým prístupom kuriéra, ktorý reaguje spontánnym postupom na prichádzajúce požiadavky zákazníkov do systému. Sú to prípady, v ktorých sa nové požiadavky objavujú až po výjazde kuriéra z logistického centra. V takejto situácii je potrebné rozhodnúť, ako najlepšie zaradiť prichádzajúce požiadavky do už naplánovanej trasy kuriéra.

Takého úlohy väčších rozmerov, ktoré sa zaoberajú optimalizovaním, resp. plánovaním dopravných sietí, patria do celočíselného programovania, a teda nájdenie ich optimálneho riešenia býva často nie len časovo náročná záležitosť, ale niekedy až nemožná. Práve preto sa k riešeniu úlohy kuriéra veľmi často používajú heuristické a metaheuristické metódy, ktoré sú schopné poskytnúť približné riešenia v uspokojivom čase. Medzi ďalšie výhody týchto metód patrí ich relatívne ľahká prispôsobivosť konkrétnym praktickým úlohám.

Cieľom príspevku je popísať a matematicky formulovať statickú úlohu kuriéra, ako jednu z modifikácií známeho problému obchodného cestujúceho. Už zo samotného názvu je zrejmé, že pôjde o priamu aplikovateľnosť tejto úlohy v oblastiach kuriérskej služby. Následne budeme prezentovať prístupy k riešeniu úlohy kuriéra pomocou rôznych algoritmov a optimalizačných techník.

## 2 OKRUŽNÉ ÚLOHY

Pod názvom kapacitné úlohy okružných jász – Capacited Vehicle Routing Problems (ďalej len CVRP) sa najčastejšie stretávame s rôznymi praktickými zovšeobecneniami úlohy obchodného cestujúceho – Travelling Salesman Problem. TSP spolu s jeho modifikáciami slúži ako základný model pre odvodenie všetkých ostatných úloh zaoberajúcich sa optimalizovaním dopravných ciest (Gutin a Punnen, 2002). Pravdepodobne prvá publikácia, v ktorej sa názov TSP objavil, bola v roku 1949 prezentovaná spisovateľkou J. B. Robinson v článku „*On the Hamiltonian game: A Traveling Salesman Problem*“. Systematické skúmanie TSP ako kombinatorického optimalizačného problému však odštartovala až práca matematikov Danzinga, Fulkersona a Johnsona z roku 1954.

Základným cieľom TSP je nájsť najkratšiu cestu obchodného cestujúceho, tzv. *Hamiltonov cyklus*, v ktorom sú všetky uzly navštívené práve raz a celková prejdená vzdialenosť je minimálna. Ide teda o hľadanie najlepšej trasy obchodníka, ktorý musí z východiskového bodu navštíviť danú množinu miest a vrátiť sa naspäť s tým, aby celkové náklady na cestu boli minimálne. Pojem „*okružné úlohy*“ je tak často spájaný s tzv. „prejazdom“ cez dané miesta, resp. s jeho „návštevou“ rôznych úsekov.

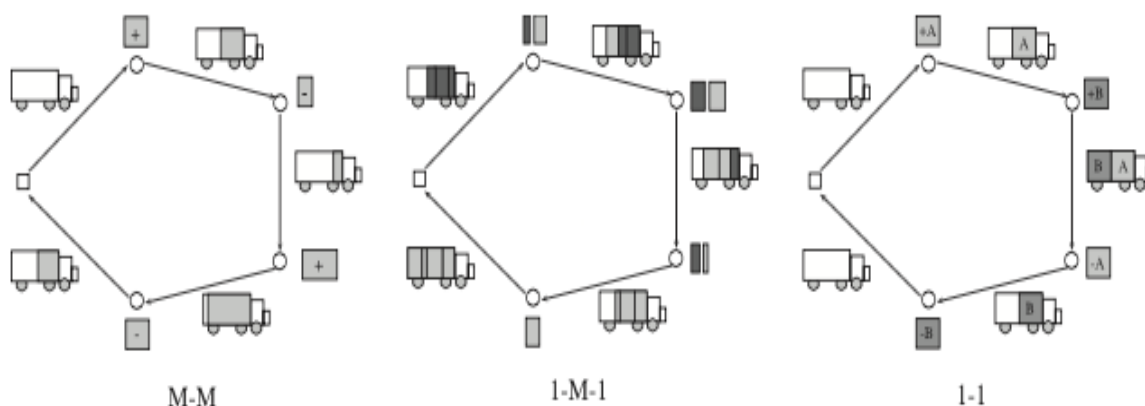
Pokiaľ ide o vymedzenie pojmu „*rozvozové úlohy*“, jedná sa nie len o rozvoz tovarov, materiálov, surovín, ľudí a pod, ale aj o ich zvoz. Jednou z najvýznamnejších publikácií, ktoré obsahujú prehľad základných variant rozvozových úloh vrátane metód ich riešenia, je zborník Totha a Viga (2002) Pod túto kategóriu zaraďujeme aj úlohu pripúšťajúcu rozvoz a zvoz v rámci jednej trasy, v zahraničnej literatúre známou ako „*Pickup and Delivery Problem*“ (ďalej len PDP), ktorej sa venovali vo svojich prácach autori ako Battarra, Cordeau, Iori (2014).

### 2.1 Pickup and Delivery Problems (PDP)

Do skupiny PDP zaraďujeme množstvo problémov, ktoré spája skutočnosť, že zadané požiadavky sú definované miestom vyzdvihnutia (*pickup*) a miestom doručenia zásielky (*delivery*), prípadne miestom vyloženia osôb (*drop off*). Týmto sa odlišujú od klasických rozvozových úloh, v ktorých sa jedná o rozvoz a zvoz tovarov z alebo do východiskového miesta a každá požiadavka je teda určená iba jedným miestom, do ktorého má požiadavka namierené.

PDP môžeme rozdeliť do troch základných skupín v závislosti od typu požiadaviek a štruktúry trasy, t.j. podľa typov miest doručenia a vyzdvihnutia. Základné kategórie PDP sú znázornené na nasledujúcom obrázku 1, pričom štvorec reprezentuje východiskový bod (sklad) a kruhy sú jednotliví zákazníci, ktorých je potrebné obslúžiť.

**Obrázok 1: Základné typy PDP**



### **Many-to many (M-M)**

Podstatou PDP typu M-M je fakt, že rozdeľuje vrcholy na množinu miest vyzdvihnutia a množinu miest doručenia tovaru. Z obrázku 1 je nám jasné, že vozidlo vychádza zo skladu prázdne a postupným prechádzaním zákazníkov, tovar buď nakladá (+), alebo vykladá (-), pričom sa nesmie presiahnuť kapacita vozidla. Typickým príkladom je „*Container Transshipment Problem*“, ktorého cieľom je prepraviť homogénny tovar od dodávateľov k odberateľom prostredníctvom distribučnej siete s obmedzenou kapacitou.

### **One to many to one (1-M-1)**

Problém je charakterizovaný existenciou tovarov, ktoré majú byť dodané zo skladu mnohým zákazníkom a ďalších tovarov, ktoré majú byť vyzdvihnuté od zákazníkov. Typickou aplikáciou tohto problému rozvoz nápojov do obchodov, pričom sa uvažuje aj so zberom prázdnych fliaš transportovaných naspäť do skladu.

### **One-to-one (1-1)**

V danom probléme uvažujeme len s tovarom s jediným miestom vyzdvihnutia a následného doručenia. Typickým príkladom je úloha kuriéra, ktorá je predmetom tohto príspevku. Požiadavka je reprezentovaná dvomi miestami, a to miestom doručenia a miestom vyzdvihnutia zásielky. Obrázok 1 vysvetľuje situáciu kuriéra, ktorý má za úlohu doručiť a vyzdvihnúť zásielky A a B. Vidíme, že najskôr zásielky naloží, a až potom ich doručí príslušnému vrcholu, resp. zákazníkovi.

## **3 ÚLOHA KURIÉRA**

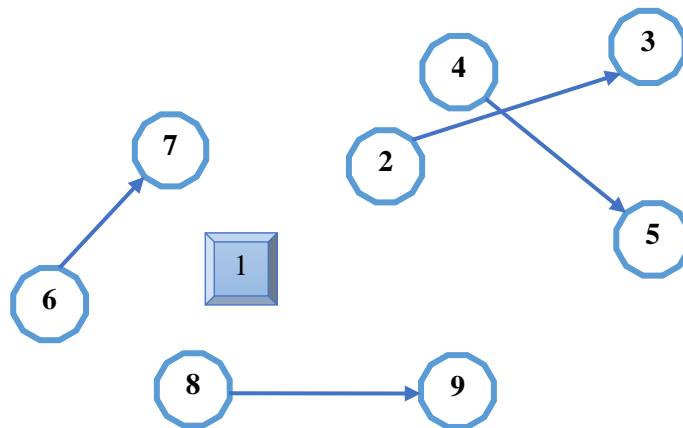
Najčastejším príkladom úlohy PDP z hľadiska jej aplikovateľnosti v praxi je úloha kuriérskej služby. Kuriéri nemusia využívať na obsluhu zákazníkov iba vozidlá, ale často využívajú aj iné dopravné prostriedky ako bicykle, ktorých kapacita je značne obmedzená. Obsluha zákazníkov spočíva v čo najrýchlejšej preprave zásielky medzi dvoma miestami, a to medzi miestom vyzdvihnutia zásielky od dodávateľa, ktorú následne doručujú zákazníkovi vo vopred dohodnutom čase (nazývame ich aj služby „od dverí k dverám“). Cieľom je minimalizovať čas doručenia s prihliadnutím na prepravné náklady.

### 3.1 Modely úloh kuriéra

Úlohu kuriéra zaraďujeme medzi modifikácie úlohy obchodného cestujúceho, pričom jej odlišnosť sa prejavuje predovšetkým v inom zedefinovaní vstupných požiadaviek a z toho vyplývajúcich dodatočných obmedzení. V úlohe kuriéra je nutné rozlišovať medzi miestom vyzdvihnutia a miestom doručenia zásielky, inak povedané, kuriér má za úlohu vyzdvihnúť zásielku u zákazníka z jedného miesta a do druhého miesta ju následne doručiť, pričom ju nie je podmienkou doručiť bezprostredne po vyzdvihnutí kuriérom. Dôležité je aj poradie navštevovania jednotlivých miest, pretože zodpovedajúce miesto doručenia nemôže byť navštívené skôr ako miesto vyzdvihnutia danej zásielky.

Na obrázku 2 ilustrujeme problém kuriéra, v ktorom sa stretávame s tromi typmi miest. Miesto, ktoré spravidla označujeme číslom 1, berieme ako východiskový bod (v našom prípade ide o logistické centrum, z ktorého kuriér, prípadne viacerí kuriéri vychádzajú). Druhým typom sú miesta vyzdvihnutia, ktoré sú reprezentované párnymi indexmi a miesta doručenia predstavujú tretí typ miest označovanými nepárnymi indexmi. Z toho nám logicky vyplýva, že do uzlu  $i+1$  nesmie kuriér doraziť skôr, ako navštívi uzol  $i$ , pričom  $i$  predstavuje nepárne číslo. V nasledujúcom obrázku nám orientované hrany znázorňujú smer doručenia zásielky jedným kuriérom.

**Obrázok 2 :** Grafické zobrazenie úlohy kuriéra



Z obrázku je nám zrejmé, že kuriér má za úlohu doručiť štyri zásielky. V našom prípade ide o uzly s nepárnymi indexmi, a preto je predovšetkým dôležité očíslovanie jednotlivých uzlov, ktoré je nutné dodržiavať vo všetkých modeloch. Cieľom kuriéra je teda navštíviť jednotlivé miesta a vrátiť sa späť do východiskového bodu, pričom hľadáme cestu s minimálnymi nákladmi.

V reálnom svete sa ale častejšie stretávame so situáciami, pri ktorých musíme rešpektovať určité ohraničenia. Najčastejšie ide o zohľadnenie časových okien, ktoré v úlohe kuriéra predstavujú jednu z najzaujímavejších rozšírení s rozsiahlym využitím v praxi. Obmedzenia sa týkajú určenia časového intervalu, v ktorom je nutné rozlišovať medzi najskôr možným začiatkom a najneskôr prípustným koncom obsluhy daného zákazníka. Inak povedané, kuriér má k dispozícii časové rozmedzie, kedy je nutné zásielku od zákazníka vyzdvihnúť a naopak, kedy ju musí zákazníkovi doručiť.

### 3.1.1 Statická úloha kuriéra

V tomto príspevku sa zameriavame na statický model kuriéra, pri ktorom na rozdiel od dynamického modelu predpokladáme, že všetky požiadavky na prepravu zásielok sú vopred známe. Inak povedané, pred zahájením optimalizácie takejto úlohy poznáme všetkých  $n$  miest vyzdvihnutia a všetkých  $n$  miesta doručenia, pričom celkovo nám je známych práve  $2n + 1$  miest (vrátane východiskového bodu). Ďalej sú nám známe náklady  $d_{ij}$ , ktoré reprezentujú vzdialenosť medzi miestami  $i$  a  $j$ . Premenná  $u_i$  reprezentuje okamžik, v ktorom kuriér dorazí do miesta  $i$ . Matematický model takejto úlohy vyzerá nasledovne (Fábry, J.,2014):

$$\text{minimalizácia } z(x) = \sum_{i=1}^{2n+1} \sum_{j=1}^{2n+1} d_{ij} x_{ij} \quad (1)$$

$$\sum_{i=1}^{2n+1} x_{ij} = 1, \quad i=1,2,\dots,2n+1, \quad (2)$$

$$\sum_{j=1}^{2n+1} x_{ij} = 1, \quad j=1,2,\dots,2n+1, \quad (3)$$

$$u_i - u_j + (2n+1)x_{ij} \leq 2n \quad i=1,2,\dots,2n+1; j=2,3,\dots,2n+1, i \neq j \quad (4)$$

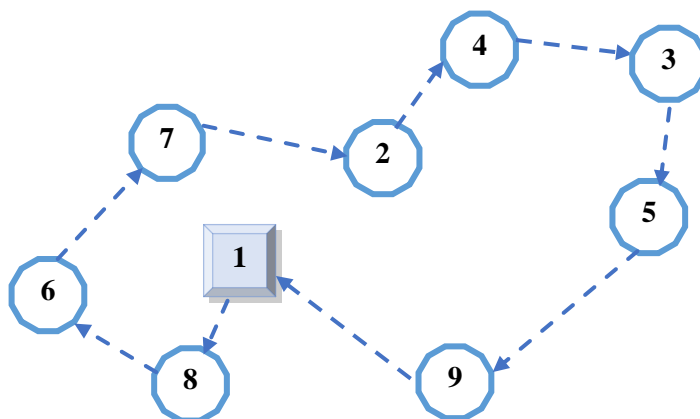
$$u_{2i} \leq u_{2i+1}, \quad i=1,2,\dots,n, \quad (5)$$

$$u_1 = 0, \quad (6)$$

$$x_{ij} \in \{0,1\}, \quad i,j=1,2,\dots,2n+1, \quad (7)$$

kde premenná  $x_{ij}$  nadobúda hodnoty 1 v prípade, ak kuriér cestuje z miesta  $i$  do miesta  $j$  a hodnoty 0 ináč. Účelová funkcia (1) reprezentuje náš cieľ, t.j. minimalizovať celkové náklady na realizáciu cesty. Sústava ohraňení (2) a (3) nám zabezpečuje, že každé miesto bude navštívené práve raz, a tým pádom zaisťuje, že každá zásielka bude vyzdvihnutá a doručená. Podmienkou (4) sa zabráňuje vytváraniu parciálnych cyklov. V úlohe jedného kuriéra nám týmto ohraňením vzniká iba jeden cyklus, pričom jeho východiskový uzol je zároveň aj jeho koncovým uzlom (uzol č. 1). Pre nás je obzvlášť dôležité ohraňenie (5), pretože rozlišuje párne a nepárne indexy jednotlivých uzlov a tým zabezpečí základnú podmienku úlohu kuriéra, že miesto doručenia nebude navštívené skôr ako miesto vyzdvihnutia zásielky.

Obrázok 3: Grafické riešenie statickej úlohy kuriéra



### 3.1.2 Statická úloha kuriéra s časovými oknami

V predchádzajúcej podkapitole sme pri hľadaní optimálnej cesty nebrali do úvahy skutočnosť, že riešenie môže byť ohraničené časovými oknami. V uvedenom modeli uvažujeme teda reálnejšiu situáciu, kedy máme dopredu určený interval vyzdvihnutia a doručenia zásielky kuriérom pri obsluhu zákazníka. V našej úlohe je teda okrem matice vzdialeností  $d_{ij}$  potrebné zohľadniť aj doby dojazdu  $u_j$  medzi všetkými miestami. Predpokladáme, že nie je možné vykonať obsluhu zákazníkov zohľadnením všetkých časových obmedzení, a preto berieme do úvahy čakanie vozidla označené premennou  $w_j$ . Potom bude v prípade nemožnosti obslúženia zákazníka v danom momente umožnené čakanie v príslušnom uzle až do najskôr prípustného začiatku obsluhy.

Účelovú funkciu teda môžeme vyjadriť ako konečný súčet času potrebného na presun a času stráveného čakaním kuriéra, pričom jej posledná časť zabezpečuje, že výsledné okamihy budú tie najmenšie možné začiatky obsluhy spĺňajúce ohraničenia týkajúce sa časových okamihov. Matematický model problému kuriéra s časovými oknami je nasledovný (Čičková Z. a kol., 2013):

$$\text{minimalizácia } z(\mathbf{x}) = \sum_{i=1}^{2n+1} \sum_{j=1}^{2n+1} d_{ij} x_{ij} + \sum_{j=1}^{2n+1} w_j + v \sum_{i=1}^{2n+1} u_i, \quad (8)$$

$$\sum_{i=1}^{2n+1} x_{ij} = 1, \quad i = 1, 2, \dots, 2n + 1, \quad (9)$$

$$\sum_{j=1}^{2n+1} x_{ij} = 1, \quad j = 1, 2, \dots, 2n + 1, \quad (10)$$

$$u_i + w_j + d_{ij} - M(1 - x_{ij}) \leq u_j, \quad i = 2n + 1, j = 1, 2, \dots, n, i \neq j \quad (11)$$

$$e_i \leq u_i \leq l_i \quad i, j = 1, 2, \dots, 2n + 1, i \neq j \quad (12)$$

$$u_{2i} \leq u_{2i+1}, \quad i = 1, 2, \dots, n, \quad (13)$$

$$u_1 = 0, \quad (14)$$

$$x_{ij} \in \{0, 1\} \quad i, j = 1, 2, \dots, 2n + 1 \quad (15)$$

$$w_j \geq 0, j = 1, 2, \dots, n, \quad (16)$$

kde premenná  $u_j$  reprezentuje okamžik, v ktorom kuriér dorazí do miesta  $i$ . Podmienky (9) a (10) nám zabezpečujú, že každé miesto bude navštívené práve raz. Nerovnicou (11) s novou premennou  $w_j$  sa vyhýbame vytváraniu podcyklov, pričom parameter  $M$  predstavuje veľmi vysokú konštantu. Zároveň slúži aj ako zábezpeka bilancie začiatkov obsluhy, t.z. reálny čas vyzdvihnutia zásielky, resp. reálny čas jej doručenia. Pre splnenie podmienok časových okien sú dôležité sústavy nerovnic (12) a (13). V sústave podmienok (12) môžeme zdefinovať nové premenné  $e_i$  a  $l_i$ , ktoré predstavujú najskôr možný začiatok a najneskôr prípustný koniec obsluhy príslušného zákazníka (v prípade vyzdvihnutia majú párny index, resp. nepárny v prípade doručenia). Ohraničenie (13) zaručuje, že konkrétna zásielka prebratá kuriérom bude doručená až po okamihu jej vyzdvihnutia a popritom zároveň dovoľuje navštívenie akékoľvek iného miesta doručenia, resp. vyzdvihnutia.

## 4 OPTIMALIZÁCIA BALÍKOVEJ PREPRAVY

Preprava zásielok po dopravných sieťach je základným procesom v dopravných systémoch a dôležitou súčasťou logistických systémov. Optimálne riadenie dopravných procesov prináša so sebou nielen lepšiu kvalitu služieb, úsporu nákladov, ale taktiež vedie k znižovaniu škodlivých vplyvov na životné prostredie. Práve tvorbou vhodných modelov docielime optimalizáciu procesov na dopravných sieťach.

Na riešenie problému obchodného cestujúceho, z ktorej vychádza naša úloha kuriéra, bolo vypracovaných viacero osobitných výpočtových postupov. Sú to nasledovné metódy (Brezina I. a kol., 2009):

- **optimalizačné metódy** - tieto metódy vedú k optimálnemu riešeniu, nie sú však vo všeobecnosti použiteľné pre úlohy s veľkým množstvom uzlov,
- **heuristické metódy** a
- **kombinované metódy**

### 4.1 Heuristické algoritmy na riešenie statickej úlohy kuriéra

Rovnako ako pri formulácii matematických modelov, tak aj pri využívaní heuristík na riešenie úlohy kuriéra je nutné použitie rôznych modifikácií. Pre úlohu kuriéra boli definované rôzne modifikácie heuristických metód najbližšieho suseda, výmen, Clarkovej-Wrightovej a vkladacej metódy. Základy spomínaných heuristík boli upravené špeciálne pre úlohu kuriéra a jej rôzne modifikácie s ohľadom na čas a kapacitu kuriéra. V tomto príspevku sa budeme venovať modifikácií algoritmu najbližšieho suseda.

#### 4.1.1 Modifikácia algoritmu najbližšieho suseda

Pri tejto metóde ide o určenie najbližšieho miesta, do ktorého je možné sa vydať z miesta, v ktorom sa kuriér práve nachádza. Na rozdiel od rozvozových úloh je nutné brať do úvahy, že kuriér nie len zásielku doručuje na rôzne miesta, ale ich aj vyzdvihuje, pričom k doručeniu nemusí dôjsť ihneď po prevzatí, pokiaľ to dovoľuje kapacita kuriéra. Modifikácia spočíva v tom, že pri vychádzaní z depa uvažujeme len miesta s párnymi indexami, v ktorých môžeme nabráť dodávku. V prípade úlohy kuriéra je teda potrebné v každom kroku algoritmu aktualizovať množinu miest pridávaním ďalších, ktoré je možné navštíviť. Po navštívení konkrétneho miesta vyzdvihnutia  $i$  je toto miesto z množiny doteraz nenavštívených miest odobraté a naopak miesto doručenia zásielky  $i+1$  do množiny pridané. Algoritmus končí vtedy, keď je množina  $A$  prázdna.

Nech  $A$  je teda množina doteraz nenavštívených miest, ktoré je v danom kroku možné navštíviť, čiže na začiatku algoritmu obsahuje iba párne čísla  $2, 4, 6$  až  $2n$ . Postupnosť navštívených miest označujeme  $S$ , vzdialenosť medzi miestami  $i$  a  $j$  je daná maticou  $d_{ij}$  a celkovú vzdialenosť, ktorú kuriér prejde označujeme premennou  $z$ . Pre úlohu jedného kuriéra môžeme algoritmus zapísať do nasledujúcich krokov (Fábry J., 2014):

*Krok 1.*  $A = \{2, 4, \dots, 2n\}, S = \{1\}, z = 0, i = 1$

*Krok 2.* Nájďme miesto  $i$  také, že  $c_{ik} = \min_{j \in A} d_{ij}$ ,

$$S = S + \{k\}, z = z + d_{ik}, A = A - \{k\},$$

ak je  $k$  párne, potom  $A = A + \{k + 1\}$ ,  
 $i = k$

Krok 3. ak je  $A = \emptyset$ , pokračujeme na krok 4, inak krok 2

Krok 4.  $S = S + \{1\}, z = z + d_{k1}$ . Koniec algoritmu

Základný proces tejto metódy opisuje krok 2. V danom riadku matice vzdialeností nájde najbližšie miesto, do ktorého je možné sa vydať. Toto miesto potom pridá do postupnosti navštívených miest a odoberá z množiny miest, ktoré je možné navštíviť. Posledná časť tohto kroku slúži na overenie, či miesto, do ktorého má kuriér namierené má páry index. Ak áno, pridá množiny do  $A$  miesto doručenia danej zásielky.

## 5 ZÁVER

Cieľom tohto príspevku bolo prezentovať modely a algoritmy na riešenie úloh kuriérskej služby. Vzhľadom na to, že úlohu kuriéra zaradujeme medzi NP-ťažké úlohy, pre riešenie rozsiahlejších a časovo náročnejších úloh možno použiť viacero metód a algoritmov, ktoré poskytujú pomerne kvalitné riešenia v akceptovateľnom čase. Jedná sa predovšetkým o heuristické algoritmy cielene vyvinuté pre úlohu kuriéra. Okrem iného sú predstavené aj optimalizačné matematické modely úlohy kuriéra v tvare celočíselného lineárneho programovania. Optimalizačné algoritmy v prípade úloh väčšieho rozmeru nemusia v reálnom čase poskytovať optimálne riešenie, a preto sa využívajú skôr heuristiky, inak povedané „približné“ postupy riešenia, ktoré síce nenachádzajú optimálne riešenie, ale keďže sú tieto postupy vždy polynomiálne, dokážu nájsť „dobrý“ výsledok v prijateľnom čase.

### Použitá literatúra

1. BATTARRA, M. CORDEAU, J.F. IORI, M. 2014. *Pickup-and-Delivery Problems for Goods Transportation In Vehicle Routing*, Bologna: Society for Industrial and Applied Mathematics and the Mathematical Optimization Society, 2014. s. 162.
2. BREZINA, I. – ČIČKOVÁ, Z. – REIFF, M. 2009. *Kvantitatívne metódy na podporu logistických procesov*, Bratislava: Vydavateľstvo EKONÓM, 2009. 56 s.
3. CENEK, P. 2003. Modelování procesů na dopravních sítích. In *AT&P journal*, 2003, č. 9. s. 98-101.
4. ČIČKOVÁ, Z. – BREZINA, I. – PEKÁR, J.: *Úloha kuriéra s časovými oknami*. In *Trendy v podnikaní 2013 [CD-ROM]*. Plzeň: Vydavateľstvo Západočeská univerzita v Plzni, 2013. 1-5 s.
5. DANTZIG, G. – FULKERSON, R., JOHNSON, S. 1954. *Solution of a large scale traveling salesman problem*. In *Journal of the Operations Research Society of America*, roč. 28, č.3, s. 497-520.
6. FÁBRY, J. 2014. Okružní a dopravní úlohy. Habilitační práce. VŠE-FIS, 2014
7. GUTIN, G. – PUNNEN, A. P. 2002. *The Traveling Salesman Problem and Its Variations*, United States: Vydavateľstvo Springer, 2002.
8. TOTH, P. – VIGO, D. 2002. *The Vehicle Routing Problem*, Montreal: Vydavateľstvo Society for Industrial and Applied Mathematics, 2002.

### Kontaktné údaje

Ing. Dana Figurová

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 948) 116 462

email: [danafigurova@gmail.com](mailto:danafigurova@gmail.com)



# VYUŽITIE PRINCÍPOV VZÁJOMNÉHO ZÁPOČTU ZÁVÄZKOV VO VYUŽÍVANÍ ENERGIE Z OBNOVITELNÝCH ZDROJOV<sup>1</sup>

## THE USE OF THE PRINCIPLES OF MUTUAL DEBTS COMPENSATION FOR USING RENEWABLE ENERGY<sup>2</sup>

*Pavel Gežík*

### **Abstrakt**

Rastúce ceny elektrickej energie a trend získavania energie z obnoviteľných zdrojov vedie k čoraz väčšiemu záujmu o získavanie elektrickej energie zo slnka. Získavanie elektrickej energie pre domácnosť môže prebiehať aj vo vlastnej réžii, čím si domácnosť znižuje spotrebu elektrickej energie odobranej zo siete. Úloha distribučnej spoločnosti je ale stále významná, keďže pre domácnosť zabezpečuje 100% pokrytie spotreby elektrickej energie. V súvislosti s výrobou elektrickej energie musia domácnosti riešiť aj problematiku prebytkov zo získanej energie. Ekonomická výhodnosť teda nespočíva len v znižovaní spotreby zo siete ale aj v možnosti ekonomického sa zbavovania prebytkov.

Pri odovzdávaní týchto prebytkov má domácnosť viacero možností, a teda aj viacero možností ako si znížiť záväzky voči distribučnej spoločnosti. Pri tomto znižovaní je vhodné využívať princípy vzájomného zápočtu záväzkov.

**Kľúčové slová:** *Obnoviteľná energia, vzájomné započítavanie záväzkov, ekonomická efektívnosť*

### **Abstract**

Rising electricity prices and the trend of renewable energy have led to an increasing interest in obtaining electricity from the sun. The households can reduce the consumption of electricity from the grid by using their own devices for obtaining renewable energy. However, the role of a distribution company is still significant, as it provides 100% coverage of electricity consumption for the households. Households also need to solve problems with surplus energy in the time of no consumption. The economic advantages are not only in the reduction of consumption from the network but also in the possibility of economic elimination of surpluses.

The households have more than one option what to do with these surpluses and therefore several options to reduce their debts to the distribution company. It is appropriate to use the principles of mutual recognition of commitments in this reduction.

**Keywords:** *Renewable energy, mutual debts compensation, economic efficiency*

---

<sup>1</sup> Tento príspevok je podporený Vedeckou grantovou agentúrou - VEGA 1/0697/15 „Optimalizačné modely a metódy ako nástroje efektívnej regulácie v modernej teórii organizácie sieťových odvetví.“

<sup>2</sup> This paper is supported by the Grant Agency of Slovak Republic - VEGA, grant no. 1/0697/15 „Optimalizačné modely a metódy ako nástroje efektívnej regulácie v modernej teórii organizácie sieťových odvetví.“

# 1 ÚVOD

Rastúce ceny elektrickej energie a trend získavania energie z obnoviteľných zdrojov vedie k čoraz väčšiemu záujmu o získavanie elektrickej energie zo slnka. Jedná sa o využívanie fotovoltaických panelov, ktoré možno inštalovať aj na majetok spotrebiteľa (strecha rodinného domu, bytového domu, pozemok, ...).

Slovensko sa tiež prispôsobuje týmto trendom a vzniká čoraz viac výrobcov energií z obnoviteľných zdrojov. Napr. v Európskej únii sa podiel takejto energie na hrubej konečnej spotrebe energie za posledných 10 rokov viac ako zdvojnásobil. Okrem získavania energií od výrobcov, ktorí využívajú tieto zdroje sa dostáva do popredia aj uvedený spôsob, kedy si spotrebiteľia vyrábajú elektrickú energiu vo vlastnej réžii.

Cieľom EÚ je dosiahnuť do roku 2020 podiel energie z obnoviteľných zdrojov na hrubej konečnej spotrebe na úrovni 20 %. Obnoviteľné zdroje energie však budú kľúčové v pokrývaní energetických potrieb EÚ aj po roku 2020. Členské štáty únie si preto stanovili nový cieľ, a to do roku 2030 dosiahnuť podiel obnoviteľných zdrojov energie na hrubej konečnej spotrebe energie na úrovni najmenej 27 %.<sup>3</sup>

Na podporenie týchto cieľov bolo vytvorených niekoľko legislatívnych smerníc EÚ ako napr. smernica 2009/28/ES, (Smernica o energii z obnoviteľných zdrojov energie). Ich vykonávaním sa však nemusia vždy dosahovať najefektívnejšie výsledky na trhu a za určitých podmienok aj štátna pomoc môže byť vhodným nástrojom, ktorý prispeje k dosahovaniu uvedených cieľov tak EÚ ale aj cieľov Slovenska v tejto oblasti. Podľa článku 4 ods. 1 smernice 2009/28/ES sú členské štáty povinné stanoviť svoje ciele pre podiel energie z obnoviteľných zdrojov v roku 2020 v týchto sektoroch aj v sektore „výroba elektriny“.

## 2 MOŽNOSTI ZÍSKAVANIA ELEKTRICKEJ ENERGIE

Získavanie elektrickej energie z obnoviteľných zdrojov najčastejšie prebieha prostredníctvom vybraného dodávateľa, ktorý elektrickú energiu nakupuje od výrobcov, ktorí využívajú obnoviteľné zdroje. Následne využívajú distribučnú spoločnosť na zabezpečenie dodávky elektrickej energie. Spotrebiteľ si následne môže zistiť, koľko ním odobranej energie bolo získaných z obnoviteľných zdrojoch.

Získavanie elektrickej energie pre domácnosť môže prebiehať aj vo vlastnej réžii spotrebiteľa, kedy zvolí využitie fotovoltaických panelov, ktoré premieňajú slnečné svetlo na jednosmerný elektrický prúd, ktorý prostredníctvom frekvenčného meniča premieňa na prúd striedavý.

Nevýhoda spomínaného získavania elektrického prúdu má veľkú nevýhodu, ktorá spočíva v skutočnosti, že spotreba elektriny v domácnosti je často vo večerných hodinách a v čase najsilnejšieho slnečného svitu má domácnosť nízku spotrebu.

Z toho dôvodu sú často k riešeniu pridávané batérie, ktoré uskladnia energiu, ktorá nebola spotrebovaná v čase jej vytvorenia. Toto riešenie je vzhľadom na cenu batérií pomerne neefektívne a cenovo nevýhodné pri premietnutí ceny batérii do ceny uskladnenej energie. Takýto spôsob získavania energie je znázornený na obrázku nižšie.

---

<sup>3</sup> [http://www.urso.gov.sk/sites/default/files/PorovnaniePodporyOZE\\_12072016.pdf](http://www.urso.gov.sk/sites/default/files/PorovnaniePodporyOZE_12072016.pdf)

*Obr. 1 - Získavanie energie z fotovoltaiických panelov a zo siete*

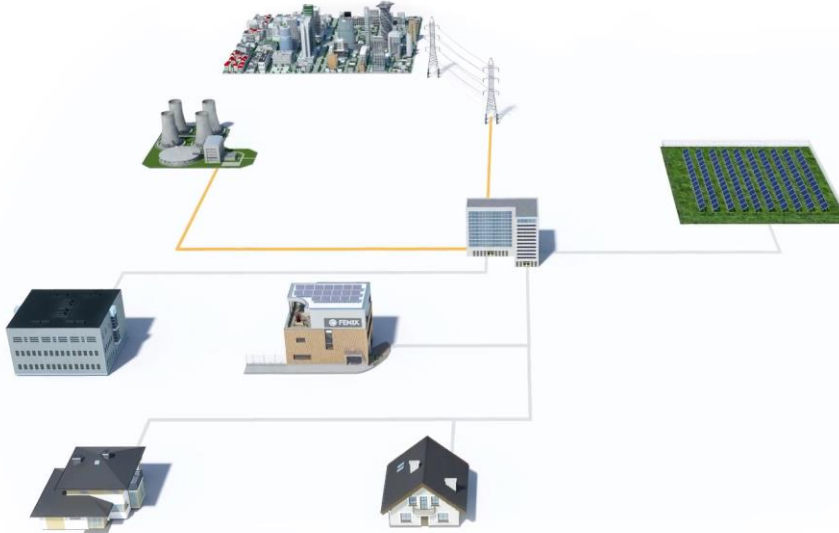


Zdroj: <http://www.fenix.sk/sk>

Na obrázku je ilustrované použitie fotovoltaiických panelov, ktoré sú napojené na batériu. Táto batéria je ale napojená aj na elektrickú sieť z dôvodu 100% dodávky elektrickej dodávky pre prevádzku domácnosti. Využívanie fotovoltaiických panelov je pre domácnosti neefektívne najmä kvôli skutočnosti, že najvyššia spotreba domácnosti je často v čase, kedy je nedostatok slnečného svitu pre výrobu potrebnej energie. Z toho dôvodu je domácnosť závislá na dodávke energie zo siete.

Tieto problémy s nesúlalom času spotreby a zisku energie z obnoviteľných zdrojov je možné riešiť aj spájaním domácností alebo všeobecne spotrebiteľov. Jedná sa o vytváranie zhlukov alebo skupín spotrebiteľov, kedy niektorí zo spotrebiteľov majú možnosť vytvárania eklektickej energie a využívajú distribútora elektrickej energie aby prerozdělil pre nich nadbytočnú energiu v čase, kedy ju nepotrebujú. Takéto riešenie je znázornené na obrázku.

*Obr. 2 – Vytváranie skupín spotrebiteľov energie a možnosti dodávky energie*



Zdroj: <http://www.fenix.sk/sk>

Podmienkou tohto riešenia je zapojenie distribútora do celého procesu. Distribútor a jeho služby sú vždy spojené s navýšením ceny. Pritom distribútor musí byť aj tak stále prítomný vo všetkých procesoch kvôli faktu, že spotrebiteľ potrebuje sústavnú dodávku.

Ideálnym riešením je vytvorenie skupiny spotrebiteľov, ktorá nevyužíva distribútora na prerozdelenie elektrickej energie ale distribúcia prebytočnej energie prebieha v rámci skupiny, kedy tí spotrebiteľia, ktorí využívajú obnoviteľné zdroje dodávajú energiu spotrebiteľom, ktorí odoberajú energiu len zo siete.

Takéto riešenie síce rieši problematiku vynechania distribútora ale zas tu vzniká problematika riešenia platieb za elektrickú energiu. Spotreba jednotlivých subjektov je rôzna a zároveň musí byť zohľadnená skutočnosť, že subjekt, ktorý vyrába energiu ju aj spotrebováva. Niektoré subjekty môžu mať systém s batériami, iné bez batérií. Spôsob úhrady platieb preto treba vždy stanoviť podľa špecifik a parametrov všetkých subjektov v skupine.

### **3 SPÔSOBY REALIZÁCIE PLATIEB ZA ENERGIU**

Základný predpoklad pre všetky spôsoby úhrady za spotrebovanú energiu je existencia distribútora elektrickej energie. Zdroje, resp. to, ako získava elektrickú energiu nie je podstatný. Samozrejme, v súvislosti s obnoviteľnými zdrojmi by bolo žiaduce aby túto energiu získal z obnoviteľných zdrojov. Ideálne od výrobcu energie, ktorý využíva na produkciu elektrického prúdu vodné elektrárne, keďže hlavný dôvod využívania distribútora spočíva v zabezpečení elektrického prúdu v čase, kedy nie je možné získať energiu zo slnečného svitu. Vodná elektráreň je vhodná náhrada a zároveň výborný doplnok, nakoľko má možnosť využívať zásobu vody v priehrade v čase reálnej spotreby. Prečerpávacie elektrárne majú ešte možnosť využívať prebytok v sieti na prečerpávanie vody, ktorú v čase nedostatku môžu využiť na tvorbu energie.

#### **3.1 Samostatný subjekt a distribútor elektrickej energie**

Najjednoduchší a zároveň základný spôsob je vzájomné započítanie záväzkov medzi subjektom a distribútorom. Z ekonomického pohľadu vzniká tak spotrebiteľovi záväzok za odobranú energiu prostredníctvom distribútora a rovnako vzniká záväzok distribútora voči spotrebiteľovi, ktorý ale pre distribútora predstavuje aj dodávateľa.

Princíp spočíva v jednoduchom odpočte hodnoty vyprodukovanej energie od spotrebovanej energie spotrebiteľom. Spotrebiteľ spotrebováva sebou vyrobenú energiu v čase jej získania a všetky prebytky odovzdáva do siete. Ekonomická výhodnosť spočíva v tom, že v čase slnečného svitu nespotrebováva energiu zo siete a odovzdaním prebytku do siete znižuje aj platbu za energiu, ktorú spotrebuje v čase, keď je fotovoltaický zdroj nepostačujúci.

Tento spôsob sa stal ale nevyhovujúci pre distribučné spoločnosti z dôvodu vyšších nárokov na vybudovanú sieť. Distribučné spoločnosti teda prestali so započítavaním, resp. s vykupovaním prebytkov od spotrebiteľov. Jedna dokonca zakázala odovzdávať prebytky do siete z dôvodu bezpečnosti, kedy nebolo možné pri servisnom zásahu garantovať bezpečnosť siete (nulové napätie) práve kvôli prebytkom v sieti.

Vychádzajúc z predchádzajúceho spôsobu, ktorý je náročné uplatniť v praxi je zřejmé, že využiteľnosť obnoviteľných zdrojov pre samostatného spotrebiteľa, resp. domácnosť má značné obmedzenia v ekonomickej výhodnosti. Malý fotovoltaický systém je nepostačujúci, resp. fixné náklady sú v tomto prípade vysoké. Väčší a veľký systém je zase neekonomický kvôli nevyužívaniu prebytkov.

Z tohto dôvodu je výhodné pre subjekty s fotovoltaickým systémom zapojenie sa do skupiny spotrebiteľov.

### 3.2 Započítavanie v rámci skupiny s využitím subjektu s veľkou spotrebou

Pri tvorbe skupín je nutné rešpektovať existujúcu sieť, resp. myslieť na isté predpoklady pri budovaní novej siete. Využitie tohto princípu ale predpokladá, že sa vytvorí skupina spotrebiteľov, ktorá bude využívať jeden subjekt na odber ich prebytkov.

V praxi by sa mohlo jednať o výrobný podnik alebo multifunkčnú budovu, resp. akýkoľvek iný subjekt, ktorý má najvyššiu spotrebu elektrickej energie v čase keď domácnosti vykazujú prebytky. Zväčša sa jedná o čas v produktívnej časti dňa, kedy je zároveň najvyššia pravdepodobnosť získania energie zo slnečného svitu. V tomto čase majú práve domácnosti nízku spotrebu vychádzajúc zo zjednodušenej skutočnosti, že väčšina členov domácnosti sa nachádza v zamestnaní.

Princíp započítavania, resp. úhrad v tomto prípade by spočíval v tom, že hlavný subjekt by odoberal elektrickú energiu od jednotlivých domácností, za ktorú by mu vznikali záväzky, ktoré by im uhrádzal a tým znižoval ich úhrady za energiu zo siete.

Tento jeden hlavný subjekt môže na seba zobrať aj úlohu distribútora elektrickej energie, kedy by on odoberal energiu zo siete a aj od domácností, ktoré sú spotrebiteľmi ale zároveň aj výrobcami energie z obnoviteľných zdrojov. V prípade, ak by mali domácnosti prebytok, tak by odoberal energiu od nich za nižšiu cenu ako zo siete v čom by spočívala ekonomická efektívnosť tohto riešenia pre hlavný subjekt. V prípade nedostatku energie v domácnostiach by domácnosti odoberali elektrickú energiu od hlavného subjektu, ktorý by ju získaval zo siete. Takýto hlavný subjekt by mal pravdepodobne nižšie ceny nákupu elektrickej energie zo siete ako domácnosti. Rozdiel v týchto cenách by predstavoval pre hlavný subjekt zisk a ekonomickú efektívnosť tohto riešenia.

Princíp započítavania pre domácnosti, ktoré sú spotrebiteľmi ale zároveň aj výrobcami energie v prípade, že by hlavný subjekt bol aj distribútorom je popísaný v časti 3.1.

### 3.3 Využitie modelu vzájomného zápočtu v skupine s viacerými druhmi spotrebiteľov

Najviac možný variant je vytvorenie skupiny medzi spotrebiteľmi s cieľom využitia prebytkov pre subjekty, ktoré nevyužívajú obnoviteľné zdroje. Takýto príklad môže predstavovať napr. ulica s rodinnými domami alebo bytový dom, kedy niektoré domácnosti investovali do zariadení na využívanie obnoviteľných zdrojov. V zjednodušenom príklade na jednej ulici so spoločnou sieťou sú na niektorých domoch fotovoltaické panely pre spotrebu týchto domov a v prípade prebytkov odovzdávajú energiu domom v ich bezprostrednom okolí bez tejto možnosti. Toto všetko v rámci jednej ulici, teda uzavretého systému so správcom. Celá ulica je podľa základného predpokladu pripojená na sieť ale distribútorovi uhrádzajú iba spotrebu nad rámec energiu, ktorú si v skupine vyprodukojú.

Dôležité je mať v rámci skupiny väčšie množstvo len spotrebiteľov aby sa zabezpečila spotreba všetkých prebytkov.

Princíp v tomto prípade môže využívať model vzájomného zápočtu (Gežik et al. 2016), a vytvorením siete medzi všetkými spotrebiteľmi je splnený predpoklad existencie úplného súvislého grafu popisujúceho možné vzťahy medzi subjektmi. Tento graf  $G = (V, H, d)$  musí byť úplný, a teda musí existovať vzťah (hrana) z množiny  $H$  medzi každým subjektom (uzlom) z množiny  $V$ . Tieto hrany sú ohodnotené dlhom medzi subjektom  $i$  a subjektom  $j$   $d_{ij}$ .

Tento dlh predstavuje odobranú energiu medzi domácnosťami. Pokiaľ neexistuje vzťah medzi subjektom  $i$  a subjektom  $j$ , tak  $d_{ij} = 0$ . Kvôli vzniku nových dlžníckych vzťahov, teda úplného grafu je nutné vytvoriť maticu s hodnotami  $c_{ij}$ ,

$$c_{ij} = \begin{cases} d_{ij} + M & \text{if exist } d_{ij} \\ 0 & \text{if } i = j \\ M & \text{if not exist } d_{ij} \end{cases} \quad (1)$$

kde  $M$  je veľké kladné reálne číslo ( $M \gg \max_{(i,j) \in H} d_{ij}$ ).

Nech  $x_{ij}$  je podiel dlhu  $c_{ij}$ , ktorý ostane po započítaní dlhov medzi subjektom  $i$  a  $j$ . Potom hodnota dlhu po započítaní je daná vzťahom

$$y_{ij} = c_{ij} x_{ij}. \quad (2)$$

Keďže  $x_{ij}$  predstavuje podiel na dlhu po započítaní (zníženého), teda pre zostatok dlhu, ktorý je prevoditeľný na percentá musí platiť

$$0 \leq x_{ij} \leq 1 \quad (3)$$

Vzájomný zápočet, teda kompenzácia dlhu predstavuje potom hodnotu  $1 - x_{ij}$ . Kompenzácia vzájomných dlhov nesmie zmeniť bilanciu dlhov subjektu  $i$  a musí platiť

$$\sum_{j=1}^n c_{ij} (1 - x_{ij}) = \sum_{j=1}^n c_{ji} (1 - x_{ji}) \quad \text{pre } i = 1, 2, \dots, n. \quad (4)$$

Následne úlohu minimalizácie dlhu je možné formulovať nasledovne:

$$\min f(x) = \sum_{i=1}^n \sum_{j=1}^n y_{ij} \quad (5)$$

$$y_{ij} = c_{ij} x_{ij} \quad (5)$$

$$\sum_{j=1}^n c_{ij} (1 - x_{ij}) - \sum_{j=1}^n c_{ji} (1 - x_{ji}) = 0 \quad \text{pre } i = 1, 2, \dots, n$$

$$0 \leq x_{ij} \leq 1$$

Výsledkom tohto modelu je potom matica platieb medzi domácnosťami, ktoré energiu vyrábajú a majú prebytky a domácnosťami, ktoré ju len spotrebovávajú

## 4 ZÁVER

Vyššie popísané možnosti sú pomerne zovšeobecnené a nereflektujú všetky špecifiká a parametre jednotlivých sietí. Podstata je kladená najmä na ekonomické riešenia a ekonomickú stránku problematiky. Technická stránka nie je podrobne popísaná a pri tvorbe teoretických skupín spotrebiteľov sa abstrahovalo od technickej náročnosti ako aj ekonomickej náročnosti na vytvorenie takejto skupiny. Vzhľadom na fakt, že sa jedná o aktuálnu problematiku je vhodné predpokladať, že bude využívaná viac v novej výstavbe, kde je možnosť vytvorenia popísaných skupín podstatne jednoduchšia.

Podstatné informácie spočívajú v popísaní možností na realizáciu úhrad, ktoré môžu napomôcť pri rozhodovaní o ekonomickej výhodnosti riešení využívajúcich obnoviteľné zdroje. Ako bolo naznačené, investícia do zariadení umožňujúcich výrobu elektrickej energie je bez využívania prebytkov ekonomicky neefektívna.

S využitím princípov započítavania je možné nie len využiť prebytky na zníženie úhrad za elektrickú energiu ale zároveň skrátenie doby návratnosti investície do uvedených zariadení. Rovnako môžu popísané možnosti ovplyvniť energetickú náročnosť bytových domov, kedy niektoré domácnosti sa rozhodnú ekonomicky podieľať na danej investícii a tak si znížiť prevádzkové náklady.

### **Použitá literatúra**

1. GEŽÍK, P., DOMONKOS, T. a BABEJ, A., 2016. *Mutual debt compensation as a minimization problem*. In Quantitative methods in economics : multiple criteria decision making XVIII : proceedings of the [18th] international scientific conference : 25th may - 27th may 2016, Vrátna, Slovakia - Bratislava : Letra Interactive, 2016. - ISBN 978-80-972328-0-1. - S. 114-118.
2. <http://www.fenix.sk/sk>
3. <http://www.urso.gov.sk>
4. [http://www.urso.gov.sk/sites/default/files/PorovnaniePodporyOZE\\_12072016.pdf](http://www.urso.gov.sk/sites/default/files/PorovnaniePodporyOZE_12072016.pdf)

### **Kontaktné údaje**

Ing. Pavel Gežík, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: +421 2 67 295 820

email: [pavel.gezik@euba.sk](mailto:pavel.gezik@euba.sk)

# QUADRATIC ESTIMATORS OF QUADRATIC VARIATION

## KVADRATICKÉ ESTIMÁTORY KVADRATICKÉ VARIACE

*Vladimír Holý*

### Abstract

An important aspect of financial high-frequency data analysis is the estimation of the quadratic variation of the price process. This is a non-trivial task as the observed price process is contaminated by the market microstructure noise making the traditional estimator called the realized variance significantly biased. However, many noise-robust estimators have been proposed in the literature. We show that the realized variance as well as two-scale estimator, realized kernel, pre-averaging estimator and least squares estimator can be formulated as quadratic form and therefore belong to the class of quadratic estimators.

**Keywords:** *High-Frequency Data, Quadratic Variation, Quadratic Estimators, Quadratic Form*

### Abstrakt

Důležitým aspektem analýzy finančních vysokofrekvenčních dat je odhad kvadratické variace cenového procesu. Jedná se o netriviální úkol, neboť pozorovaný cenový proces je kontaminován mikrostrukturním šumem, který způsobuje významné vychýlení v tradičním estimátoru nazvaném realizovaná variace. V odborné literatuře ovšem bylo navrženo mnoho estimátorů robustních k šumu. Ukážeme, že realizovanou varianci, stejně jako two-scale estimátor, realizovaný kernel, pre-averaging estimátor a estimátor minimálních čtverců lze formulovat jako kvadratickou formu, a proto patří do třídy kvadratických odhadů.

**Keywords:** *Vysokofrekvenční data, Kvadratická variace, Kvadratické estimátory, Kvadratická forma*

## 1 INTRODUCTION

Volatility of a financial high-frequency process is often measured by the *quadratic variation*. The quadratic variation of the process  $P_t$  is defined as

$$QV_{(0,1)} = \text{plim}_{\Delta^N \rightarrow 0} \sum_{i=2}^N \left( P_{T_i} - P_{T_{i-1}} \right)^2, \quad (1)$$

where  $\text{plim}$  denotes the limit in probability and  $\Delta^N = \max\{T_2 - T_1, T_3 - T_2, \dots, T_N - T_{N-1}\}$  is the maximal lag between the observations. In practice, we do not observe continuous price process  $P_t$ . Instead, we observe discrete process

$$X_i = P_{T_i} + E_i, \quad (2)$$

where  $P_t$  is the *efficient price* with continuous time  $t$  observed at discrete times  $T_1, \dots, T_N$  and  $E_i$  is the *market microstructure noise*. Furthermore, we denote the return process as  $Y_{j,i} = X_i - X_j$ . The market microstructure noise is caused by bid-ask spread, discreteness of price values, discreteness of times of price changes and various informational effects. At higher frequencies, it causes a significant bias in realized variance, a traditional estimator of the quadratic variation.



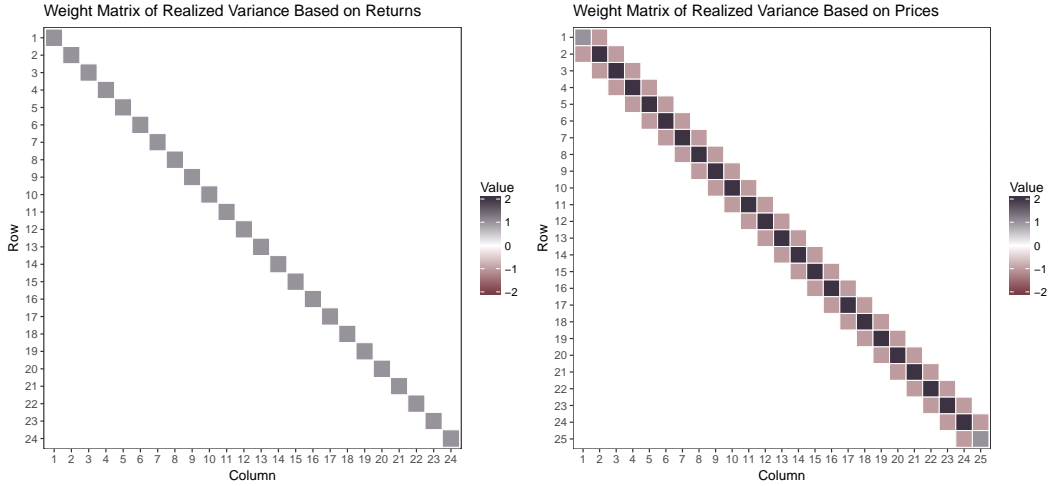


Figure 1: Quadratic form of realized variance with  $N = 25$ .

However, there are many estimators robust to the noise such as *two-scale estimator* [6], *realized kernel* [2], *pre-averaging estimator* [3] and *least squares estimator* [4]. The class of *quadratic estimators* was introduced by [5] and independently by [1]. Estimators in this class can be formulated as quadratic form, i.e.

$$QE_{(0,1)}^N = \sum_{i=2}^N \sum_{j=2}^N Y_{i-1,i} q_{i,j} Y_{j-1,j} = \mathbf{Y}' \mathbf{Q} \mathbf{Y}, \quad (3)$$

where  $\mathbf{Q}$  is a matrix of weights determining an estimator. It can also be rewritten using the actual prices instead of returns as

$$QE_{(0,1)}^N = \mathbf{Y}' \mathbf{Q} \mathbf{Y} = \mathbf{X}' \mathbf{T}' \mathbf{Q} \mathbf{T} \mathbf{X} = \mathbf{X}' \mathbf{R} \mathbf{X}, \quad (4)$$

where  $\mathbf{T}$  is a matrix with  $N - 1$  rows and  $N$  columns given by elements

$$t_{i,j} = \begin{cases} 1 & \text{for } j = i + 1, \\ -1 & \text{for } j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

## 2 REALIZED VARIANCE

The most simple estimator of quadratic variation is the *realized variance* defined as

$$RV_{(0,1)}^N = \sum_{i=2}^N (X_i - X_{i-1})^2 = \sum_{i=2}^N Y_{i-1,i}^2. \quad (6)$$

It can easily be expressed as quadratic estimator using the weight matrix  $\mathbf{Q}_{RV}^N$  given by elements

$$q_{i,j} = \begin{cases} 1 & \text{for } j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

An example of this weight matrix is shown in Figure 1. Realized variance is biased in the presence of market microstructure noise  $E_i$ . However, it is a cornerstone for more sophisticated noise-robust methods. The bias of realized variance can be reduced by sampling at lower frequencies.

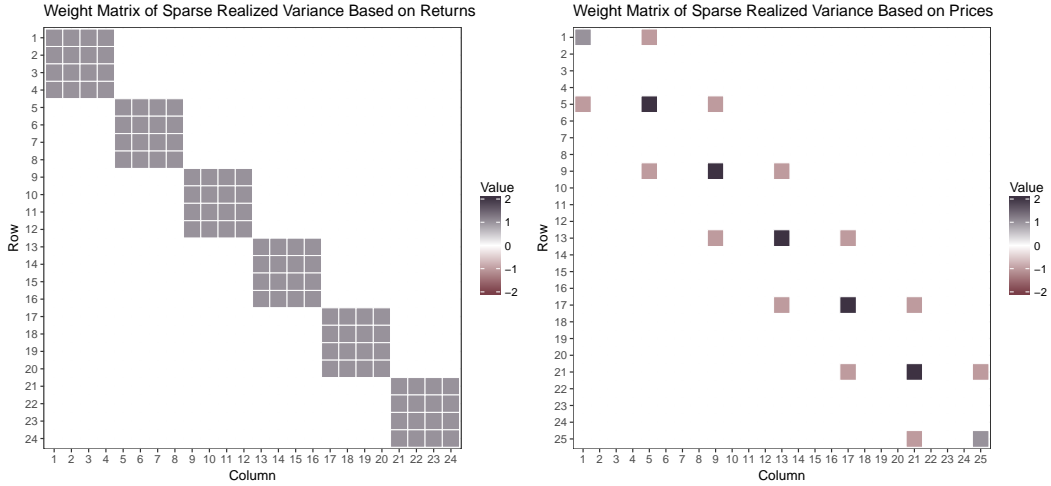


Figure 2: Quadratic form of sparse realized variance with  $N = 25$ ,  $h = 1$  and  $s = 4$ .

This approach is called the *sparse realized variance*. Let  $h$  denote the initial observation and  $s$  denote the sampling frequency of ticks. For example  $h = 2$  and  $s = 3$  would correspond to observations at times  $\{T_2, T_5, T_8, T_{11}, \dots\}$ . The number of used observations is then  $N_{h,s} = \lfloor \frac{N-h}{s} \rfloor$ . The sparse realized variance is defined as

$$SRV_{(0,1)}^{N,h,s} = \sum_{i=1}^{N_{h,s}} (X_{is+h} - X_{(i-1)s+h})^2 = \sum_{i=1}^{N_{h,s}} Y_{(i-1)s+h, is+h}^2 \quad (8)$$

Its weight matrix in quadratic representation  $\mathcal{Q}_{SRV}^{N,h,s}$  is given by elements

$$q_{i,j} = \begin{cases} 1 & \text{for } (k-1)s+h \leq i, j \leq ks+h-1, k=1, \dots, N_{h,s}, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

It is visualized in Figure 2. Sparse realized variance uses only a fraction of available observations. To fully utilize all data, the *average realized variance* can be adopted. It averages sparse realized variances over subgrids given by different initial observations  $h$ . It is defined as

$$ARV_{(0,1)}^{N,s} = \frac{1}{s} \sum_{h=1}^s RV_{(0,1)}^{N,h,s}. \quad (10)$$

The weight matrix of average realized variance is shown in Figure 3 and is given by

$$\mathcal{Q}_{ARV}^{N,s} = \frac{1}{s} \sum_{h=1}^s \mathcal{Q}_{RV}^{N,h,s}. \quad (11)$$

Although this approach reduces the impact of the noise, the average realized variance is still a biased estimator of the quadratic variation.

### 3 TWO-SCALE ESTIMATOR

The first unbiased and consistent estimator of the quadratic variation is the *two-scale estimator* [6]. It combines average realized variance at lower frequency as a biased estimate of quadratic



Figure 3: Quadratic form of average realized variance with  $N = 25$  and  $s = 4$ .

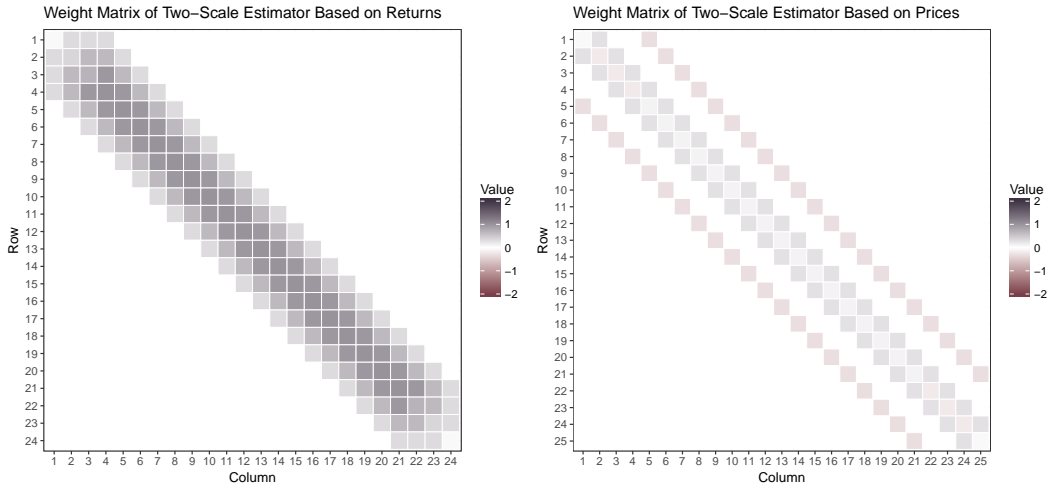


Figure 4: Quadratic form of two-scale estimator with  $N = 25$  and  $s = 4$ .

variation with realized variance at highest possible frequency as an estimator of noise variance (and therefore the bias). It is defined as

$$TSE_{(0,1)}^{N,s} = \left(1 - \frac{\tilde{N}_s}{N}\right)^{-1} \left(ARV_{(0,1)}^{N,s} - \frac{\tilde{N}_s}{N}RV_{(0,1)}^N\right), \quad (12)$$

where  $\tilde{N}_s = \frac{1}{s} \sum_{h=1}^s N_{h,s}$ . Two-scale estimator is a quadratic estimator with weight matrix

$$\mathbf{Q}_{TSE}^{N,s} = \left(1 - \frac{\tilde{N}_s}{N}\right)^{-1} \mathbf{Q}_{ARV}^{N,s} - \left(1 - \frac{\tilde{N}_s}{N}\right)^{-1} \frac{\tilde{N}_s}{N} \mathbf{Q}_{RV}^N. \quad (13)$$

As we can see in Figure 4, the structure is similar to average realized variance. However, unlike the average realized variance it is a consistent estimator.

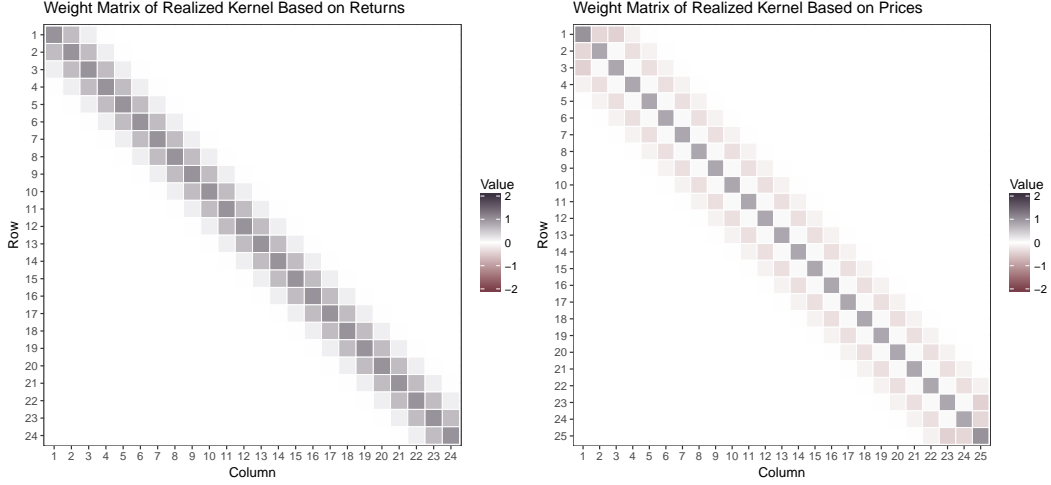


Figure 5: Quadratic form of realized kernel with  $N = 25$  and  $L = 4$ .

## 4 REALIZED KERNEL

A very popular consistent estimator is the *realized kernel* [2] defined as

$$RK_{(0,1)}^{N,L} = RV_{(0,1)}^n + \sum_{l=1}^L K\left(\frac{l-1}{L}\right) \left(RA_{(0,1)}^{N,l} + RA_{(0,1)}^{N,-l}\right), \quad (14)$$

where  $RA_{(0,1)}^{N,l}$  is the *realized autocovariance* defined as

$$RA_{(0,1)}^{N,l} = \sum_{i=\max(2,l+2)}^{\min(N,N+l)} (X_i - X_{i-1})(X_{i-l} - X_{i-l-1}) = \sum_{i=\max(2,l+2)}^{\min(N,N+l)} Y_{i-1,i} Y_{i-l-1,i-l} \quad (15)$$

and  $K(\cdot)$  is a kernel function. We use the Tukey-Hanning kernel of order two given by

$$K(x) = \sin^2\left(\frac{\pi}{2}(1-x)^2\right). \quad (16)$$

Realized kernel can be expressed in quadratic form with weight matrix  $\mathcal{Q}_{RK}^{N,L}$  given by elements

$$q_{i,j} = \begin{cases} 1 & \text{for } j = i, \\ K\left(\frac{l-1}{L}\right) & \text{for } |i-j| = l, l = 1, \dots, L, \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

An example of the structure of the weight matrix is shown in Figure 5.

## 5 PRE-AVERAGING ESTIMATOR

The market microstructure noise can be removed by locally averaging returns. The *pre-averaging estimator* [3] is based on this idea. Averaged returns are defined as

$$Z_{i,i-1} = \sum_{l=1}^{k_N} G\left(\frac{l}{k_N}\right) Y_{i+l-1,i+l-2}, \quad (18)$$

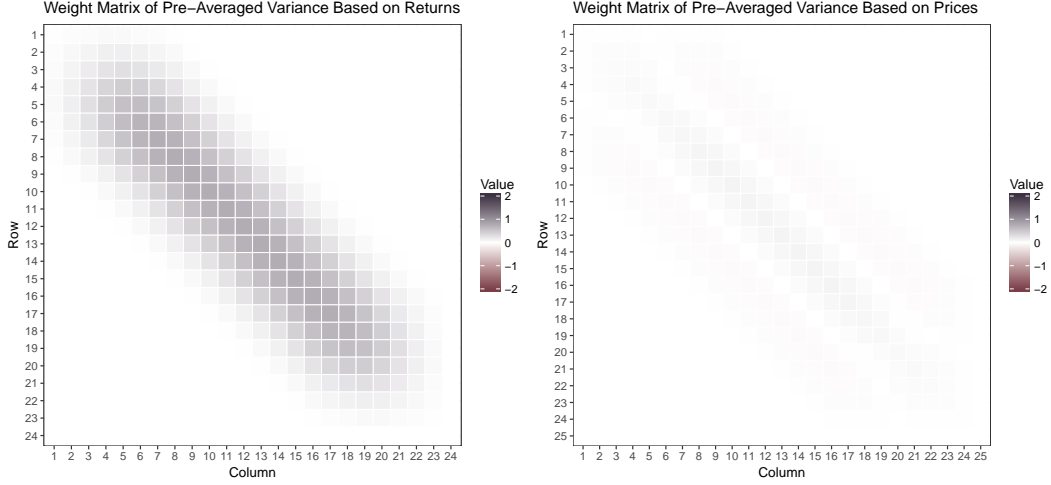


Figure 6: Quadratic form of pre-averaged variance with  $N = 25$ ,  $\theta = 0.8$  and  $k_N = 9$ .

where  $G(\cdot)$  is a function given by

$$G(x) = \min(x, 1 - x). \quad (19)$$

A direct analogue of realized variance with averaged returns is given by

$$PAV_{(0,1)}^{N,\theta,k_N} = \sum_{i=2}^{N-k_N} Z_{i-1,i}^2. \quad (20)$$

This can be formulated as quadratic form with weight matrix  $\mathbf{Q}_{PAV}^{N,\theta,k_N} = \mathbf{S}'\mathbf{S}$ , where  $\mathbf{S}$  is a matrix with  $N - l$  rows and  $N - 1$  columns given by elements

$$s_{i,j} = \begin{cases} G\left(\frac{l}{k_N}\right) & \text{for } j = i + l - 1, l = 1, \dots, k_N, \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

The structure of the weight matrix is shown in Figure 6. Pre-averaging estimator is then defined as

$$PAE_{(0,1)}^{N,\theta,k_N} = \left(1 - \frac{\psi_1^{k_N} N^{-1}}{2\theta^2 \psi_2^{k_N}}\right)^{-1} \left( \frac{\sqrt{N}}{(N - k_N + 2)\theta \psi_2^{k_N}} PAV_{(0,1)}^{N,\theta,k_N} - \frac{\psi_1^{k_N} N^{-1}}{2\theta^2 \psi_2^{k_N}} RV_{(0,1)}^N \right), \quad (22)$$

where

$$\begin{aligned} \psi_1^{k_N} &= k_N \sum_{l=1}^{k_N} \left( G\left(\frac{l+1}{k_N}\right) - G\left(\frac{l}{k_N}\right) \right)^2, \\ \psi_2^{k_N} &= \frac{1}{k_N} \sum_{l=1}^{k_N-1} G^2\left(\frac{l}{k_N}\right). \end{aligned} \quad (23)$$

It can also be formulated as quadratic form with weight matrix

$$\mathbf{Q}_{PAE}^{N,\theta,k_N} = \frac{2\theta\sqrt{N}}{(2\theta^2 - \psi_1^{k_N} N^{-1})(N - k_n + 2)} \mathbf{Q}_{PAV}^{N,\theta,k_N} - \frac{\psi_1^{k_N}}{2\theta^2 \psi_2^{k_N} N - \psi_1^{k_N}} \mathbf{Q}_{RV}^N. \quad (24)$$

The weight matrix of pre-averaging estimator is visualized in Figure 7.

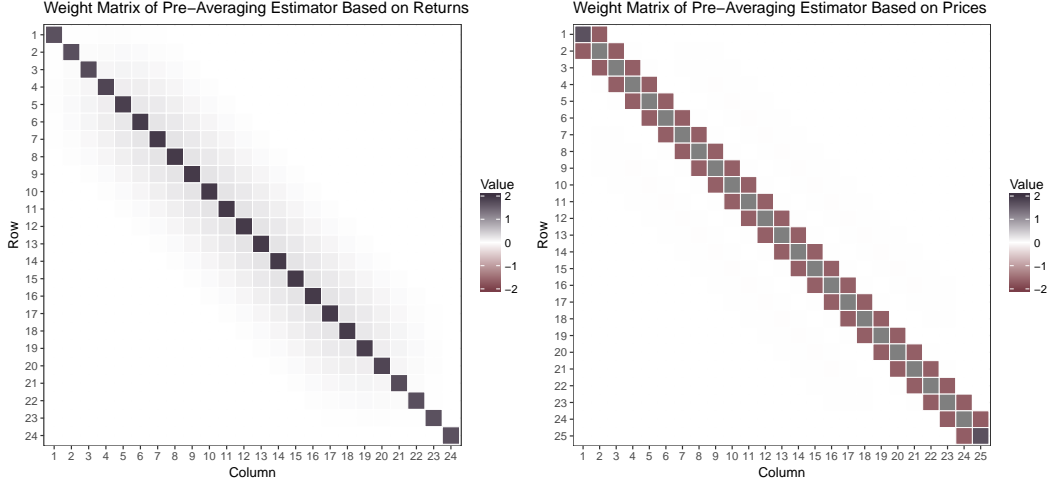


Figure 7: Quadratic form of pre-averaging estimator with  $N = 25$ ,  $\theta = 0.8$  and  $k_N = 9$ .

## 6 LEAST SQUARES ESTIMATOR

The *least squares estimator* [4] utilizes linear regression to capture the bias of realized variance for different frequencies. The expected value of sparse realized variance is given by

$$E[SRV_{(0,1)}^{N,h,s}] = QV_{(0,1)} + 2N_{h,s}\omega^2, \quad s = 1, \dots, S, \quad h = 1, \dots, s. \quad (25)$$

To find values of quadratic variation  $QV_{(0,1)}$  and noise variance  $\omega^2$ , the least squares are minimized resulting in the quadratic variation estimator

$$LSE_{(0,1)}^{N,S} = \frac{\bar{N}_2 \sum_{h=1}^s SRV_{(0,1)}^{N,h,s} - \bar{N}_1 \sum_{h=1}^s N_{h,s} SRV_{(0,1)}^{N,h,s}}{\bar{N}_0 \bar{N}_2 - \bar{N}_1^2}, \quad (26)$$

where  $\bar{N}_d = \sum_{s=1}^S \sum_{h=1}^s N_{h,s}^d$ . This is a quadratic estimator with weight matrix

$$\mathbf{Q}_{LSE}^{N,S} = \sum_{s=1}^S \sum_{h=1}^s \frac{\bar{N}_2 - \bar{N}_1 N_{h,s}}{\bar{N}_0 \bar{N}_2 - \bar{N}_1^2} \mathbf{Q}_{SRV}^{N,h,s}. \quad (27)$$

The structure of the weight matrix of the least squares estimator is similar to the two-scale estimator as we can see in Figure 8. This is because both methods are based on sparse realized variances.

## 7 CONCLUSION

We analyze several estimators of quadratic variance. We show that the realized variance as well as noise-robust methods such as two-scale estimator, realized kernel, pre-averaging estimator and least squares estimator are all quadratic estimators. This means they can be formulated as quadratic form. We derive the weight matrices of quadratic forms of the estimators and examine their structure using visualizations.

### Acknowledgements

This work was supported by the Internal Grant Agency of University of Economics, Prague under Grant F4/63/2016.

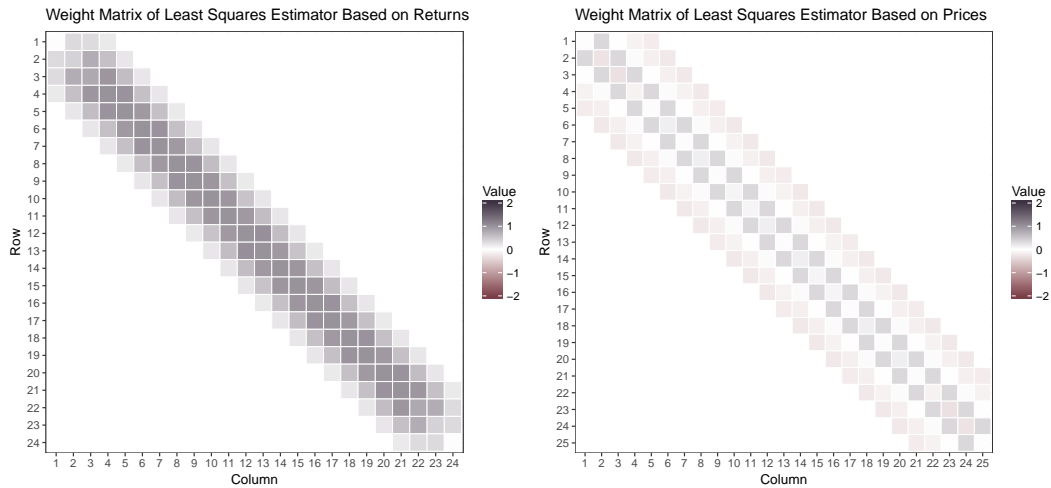


Figure 8: Quadratic form of least squares estimator with  $N = 25$  and  $S = 4$ .

## References

- [1] ANDERSEN, T. G., BOLLERSLEV, T. a MEDDAHI, N. (2011). Realized Volatility Forecasting and Market Microstructure Noise. *Journal of Econometrics*, **160**(1), 220–234. ISSN 0304-4076. doi: 10.1016/j.jeconom.2010.03.032.
- [2] BARNDORFF-NIELSEN, O. E., HANSEN, P. R., LUNDE, A. a SHEPHARD, N. (2008). Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise. *Econometrica*, **76**(6), 1481–1536. ISSN 0012-9682. doi: 10.3982/ecta6495.
- [3] JACOD, J., LI, Y., MYKLAND, P. A., PODOLSKIJ, M. a VETTER, M. (2009). Microstructure Noise in the Continuous Case: The Pre-Averaging Approach. *Stochastic Processes and their Applications*, **119**(7), 2249–2276. ISSN 0304-4149. doi: 10.1016/j.spa.2008.11.004.
- [4] NOLTE, I. a VOEV, V. (2012). Least Squares Inference on Integrated Volatility and the Relationship Between Efficient Prices and Noise. *Journal of Business & Economic Statistics*, **30**(1), 94–108. ISSN 0735-0015. doi: 10.1080/10473289.2011.637876.
- [5] SUN, Y. (2006). Best Quadratic Unbiased Estimators of Integrated Variance in the Presence of Market Microstructure Noise. URL <https://ssrn.com/abstract=1714751>.
- [6] ZHANG, L., MYKLAND, P. A. a AÏT-SAHALIA, Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, **100**(472), 1394–1411. ISSN 0162-1459. doi: 10.2307/27590680.

## Contact

Vladimír Holý  
 University of Economics, Prague  
 Winston Churchill Square 1938/4, 130 67 Prague 3, Czech Republic  
 E-mail: vladimir.holy@vse.cz

# MODELOVANIE OPTIMÁLNEJ ŠTRUKTÚRY FINANCOVANIA PODNIKU PRI NEÚPLNEJ INFORMÁCII

## MODELING OPTIMAL STRUCTURE OF FINANCING COMPANY WITH INCOMPLETE INFORMATION

*Eduard Hozlár*

### **Abstrakt**

Príspevok je venovaný konštrukcii optimálnej štruktúry financovania výroby pri náhodnom dopyte na produkciu. Vychádza sa z popisu a vlastností zdrojov financovania podnikových aktivít. Formuluje sa základný deterministický model a jeho rozšírenie o stochastický dopyt. Analyzuje sa dopyt s rovnomerným rozdelením pravdepodobností. Nakoniec je formulovaný dvojetapový stochastický model financovania s úverom, ktorý zohľadňuje optimálnu kompenzáciu odchýlok vypočítaného výrobného programu od realizovaného dopytu.

***Kľúčové slová:** stochastický model, financovanie, optimalizácia*

### **Abstract**

The article describes the creation optimal structure of production financing at random demand for products. It is based on the description and characteristics of the financing of the company's production activities as well as on the description financing of the sources. The basic deterministic model and its extension with stochastic demand are formulated. The demand with a uniform probability distribution is analyzed. Finally, a two-stage stochastic model of production financing with debt is formulated. The model takes into account the optimal compensation of the deviations of the calculated production program from the realized demand.

***Keywords:** stochastic model, financing, optimization*

## **1. ŠTRUKTÚRA FINANCOVANIA PODNIKU**

Budeme predpokladať, že finančnú štruktúru podniku tvorí podiel jednotlivých zdrojov financovania na celkových finančných zdrojoch podniku. Kapitálová štruktúra podniku je súčasť finančnej štruktúry, ktorá je spojená s dlhodobými zdrojmi financovania. Pri skúmaní hrá dôležitú úlohu statické hľadisko finančnej štruktúry (k určitému dátumu) ako aj dynamické hľadisko, kedy sa skúma hlavne zmena kapitálovej štruktúry v čase.

Najjednoduchší klasický prístup ku skúmaniu spočíva v tom, že zmena finančnej štruktúry sa prejavuje v trhovej hodnote podniku. Vychádza sa z toho, že finančný leverage (finančná páka) znižuje priemerné náklady na podnikový kapitál (t.j. vážený aritmetický priemer nákladov vlastného a cudzieho kapitálu) a tým zvyšuje trhovú hodnotu podniku. Finančnej štruktúre podniku sa venuje napr. Vlachynský (2009), Brealey, Myers, Allen, (2014).

Vážené priemerné náklady na celkový kapitál  $r_A$  (v literatúre označovaný ako WACC - Weighted Average Cost of Capital) predstavujú priemernú mieru výnosu, ktorú požadujú



poskytovatelia vlastného kapitálu (majitelia spoločnosti, investori) a poskytovatelia dlhového kapitálu (banky, majitelia dlhopisov apod.). Preto sa odporúča pri splnení aplikačných podmienok použiť  $r_A$  pre diskontovanie peňažných tokov pri oceňovaniach. Na výpočet priemerných nákladov kapitálu a ich použitie pre stanovovanie tržnej ceny podniku exustuje viacero prístupov. Jeden z nich je popísaný na *cfo.sk*. Výpočet  $r_A$  má nasledovné parametre:

$$r_A = r_E \frac{E}{E + D} + r_D \frac{D}{E + D} \quad (1)$$

kde

$r_E$  – náklady na vlastný kapitál, t.j. výnos, ktorý požadujú majitelia spoločnosti (v podobe dividend a zvýšenia hodnoty spoločnosti),

$r_D$  – náklady na dlhový kapitál po zdanení, t.j. výnos, ktorý požadujú poskytovatelia dlhu (napr. úrok požadovaný bankou), očistený o úrokový daňový štít,

$E$  – trhovú hodnotu vlastného kapitálu (imania) spoločnosti,

$D$  – trhovú hodnotu dlhového kapitálu spoločnosti, pričom  $E$  a  $D$  slúžia ako váhy pre výpočet váženého priemeru.

### 1.1. Náklady vlastného kapitálu

Náklady na vlastný kapitál  $r_E$  predstavujú výnos, ktorý z konkrétnej investície očakávajú majitelia / akcionári napríklad v podobe dividend, alebo v podobe zvýšenia hodnoty svojich podielov. Pre výpočet  $r_E$  sa najčastejšie používa model CAPM (Capital Asset Pricing Model). Základný vzorec pre výpočet  $r_E$  je nasledovný:

$$r_E = r_f + r_p \beta + r_s$$

$r_f$  – výnos z investície do bezrizikových aktív, ktorý investorom prinášajú „bezrizikové“ aktíva. Ak firma podniká v prevažnej miere na Slovensku a jej obchodný styk prebieha v eurách, môže použiť výnosy slovenských štátnych dlhopisov.

$r_p$  – trhovú riziková prirážka, t.j. dodatočný výnos nad úroveň výnosu z bezrizikovej investície, ktorý požadujú investori do (viac rizikového) akciového trhu. Tento rozdiel je potom potrebné upraviť o riziko krajiny.

$\beta$  – beta koeficient, ktorý vyjadruje rizikovosť konkrétneho odvetvia alebo spoločnosti v pomere k riziku z investície do celkového akciového trhu, ktorý má koeficient beta rovný 1. V modeli CAPM je do vzorca potrebné dosadiť betu daného odvetvia.

$r_s$  – špecifická riziková prirážka týkajúca sa konkrétnej spoločnosti (napr. prirážka za veľkosť spoločnosti)

### 1.2. Náklady cudzieho kapitálu

Získavanie cudzieho kapitálu podniku je spojené s dlhom, ktorý podnik musí v určitej lehote a výške splatiť.

Pri nákladoch na dlhový kapitál  $r_D$  sa berie do úvahy daňový úrokový štít, t.j. používajú sa náklady na dlhový kapitál po zdanení:

$$r_D = i(1 - d)$$

kde  $i$  je úroková sadzba dlhového kapitálu a  $d$  je sadzba dane z príjmov.

### 1.3. Špecifikácia matematického modelu financovania podniku

V ďalšom budeme uvažovať výrobný podnik, ktorý na svoju výrobnú činnosť využíva vlastné aj cudzie zdroje financovania.

Uvažujme nasledovný model financovania výroby v tvare

Minimalizovať finančné náklady

$$N = \sum_{j=1}^n c_j x_j \quad (2)$$

za podmienok

$$\sum_{j=1}^n z_j x_j \geq Z \quad (3)$$

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, m \quad (4)$$

$$x_j = t_j(\omega), \quad j = 1, 2, \dots, n \quad (5)$$

kde

$c_j$  – jednotkové celkové náklady  $j$ -teho výrobku

$z_j$  - ziskovosť  $j$ -teho výrobku

$a_{ij}$  – materiálové náklady na  $i$ -ty zdroj, ktorý sa spotrebuje na výrobu jednej jednotky  $j$ -teho výrobku

$b_i$  – disponibilné množstvo  $i$ -teho zdroja (materiál, pracovná sila,...)

$x_j$  – množstvo výroby  $j$ -teho výrobku

$t_j(\omega)$  – nezáporný náhodný dopyt  $j$ -teho výrobku pri realizácii náhody  $\omega$

$N$  - celkové finančné náklady na výrobu výrobkov

$Z$  – požadovaný zisk

Nech  $F_v$  je disponibilné množstvo vlastných finančných zdrojov na krytie celkových nákladov výroby a nech  $F_c$  je požadované množstvo úverových zdrojov. Je zrejmé, že musí platiť

$$F_v + F_c \geq N. \quad (6)$$

Vzhľadom na rovnaké financovanie potrieb vlastnými a cudzími zdrojmi s cieľom vytvoriť čo najvyššiu hodnotu zavedieme nasledovnú proporciu :

$$F_c = y F_v, \quad (7)$$

kde

$$0 \leq y \leq 2 \quad (8)$$

t.j. predpokladáme, že cudzie zdroje nemôžu byť väčšie ako dvojnásobok vlastných zdrojov.

Potom z (1) a (5) dostaneme

$$\sum_{j=1}^n c_j x_j - y F_v \leq F_v \quad (9)$$

*Poznámka:* Z uvedeného vyplýva, že ak  $r_A$  sú priemerné náklady na celkové finančné zdroje, potom vzťah medzi nákladmi na vlastné a cudzie zdroje podľa (1) a (7) je

$$r_E = r_A + (r_A - r_D)y.$$

To znamená, že ak cudzie zdroje nebudeme mať ( $y = 0$ ), priemerné náklady celkových zdrojov sa rovnajú nákladom vlastných zdrojov, ktoré sú pre podnik najcennejšie.

Ak  $F_v : F_c = 1:1$ , t.j.  $y = 1$ , potom  $r_E = 2r_A - r_D$ .

## 2. STOCHASTICKÝ MODEL FINANCOVANIA PODNIKU

Nech náhodný dopyt po rôznych výrobkoch je nezávislý. Potom v prípade znalosti pravdepodobnostného rozdelenia môžeme dopytové ohraničenia formulovať buď štatisticky alebo pravdepodobnostne.

V ďalšom budeme predpokladať, že dopyt  $t_j(\omega)$  má rovnomerné rozdelenie v intervale  $< d_j, h_j >$ . Potom odbytové ohraničenie, v prípade požadovania jeho splnenia v priemere, by bolo deterministické

$$x_j = 1/2 (d_j + h_j), \quad j = 1, 2, \dots, n \quad (10)$$

V prípade pravdepodobnostného ohraničenia distribučná funkcia pre  $t_j(\omega)$  má tvar

$$\begin{aligned} F_j(y) &= 0 \text{ pre } y \leq d_j \\ F_j(y) &= \frac{y - d_j}{h_j - d_j} \text{ pre } d_j \leq y \leq h_j \\ F_j(y) &= 1 \text{ pre } y \geq h_j. \end{aligned} \quad (11)$$

Ohraničenie

$$P(x_j \geq t_j(\omega)) \geq \alpha_j$$

možno transformovať na deterministický ekvivalent

$$x_j \geq d_j + \alpha_j (h_j - d_j) \quad (12)$$

Riešenie deterministických úloh (2) až (4), (8), (9), v ktorých sa ohraničenie na dopyt (5) zamení ohraničením (10) alebo (12) vyvoláva rôzne nároky na financovanie výroby. Z celkových minimálnych nákladov  $N$  možno prostredníctvom vypočítanej hodnoty  $y$  a zadanej hodnoty vlastných zdrojov vypočítať veľkosť cudzích (úverových) zdrojov.

Ak teda  $r_D$  označuje náklady na skutočne čerpané množstvo *úverového* zdroja financovania, potom finančné náklady znižujú zisk o  $r_D F_c = r_D y F_v$ .

Okrem toho, zisk z výrobného programu znižujú aj náklady (penále, resp. iné náklady, napr. skladovania) pri nedodržaní požadovaných objemov dodávok, ktoré vzniknú pri realizácii dopytu  $t_j(\omega)$ . Penalizačná funkcia  $Q$  súvisiaca s odchýlkami skutočnej dodávky od požadovaného objemu má tiež náhodný charakter a má tvar

$$Q(\mathbf{x}, \mathbf{t}(\omega)) = \sum_{d_j(\omega) - x_j \geq 0} q_j^+(t_j(\omega) - x_j) - \sum_{d_j(\omega) - x_j \leq 0} q_j^-(t_j(\omega) - x_j) \quad (13)$$

kde

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \quad \mathbf{t}(\omega) = (t_1(\omega), t_2(\omega), \dots, t_n(\omega))$$

$q_j^+$  - je penále za stratu, ktorá vznikne pri väčšom dopyte ako je ponuka  $x_j$

$q_j^-$  - je penále za stratu, ktorá vznikne pri menšom dopyte ako je ponuka  $x_j$ .

Úlohou nájsť vektor  $\mathbf{x}$ , pri ktorom sa dosahuje minimum priemerných penalizačných nákladov. To sú vlastne indukované ohraničenia pre výber riešenia  $\mathbf{x}$ .

## 2.1 Dvojetapový stochastický model financovania výroby

V uvedenom postupe modelovania sme ukázali, že riešením deterministickej úlohy (2) až (4), (8), (9), keď ignorujeme náhodný dopyt, dostaneme taký výrobný program, ktorý zabezpečuje požadované hodnoty zisku pri minimálnych nákladoch bez ohľadu ich finančného krytia.

Následne riešime otázku finančného krytia takéhoto predbežného plánu výroby. Prípadné náklady financovania znižujú vytvorený zisk.

Ďalej skúmame, či plánované množstvá výroby pokryjú realizovaný dopyt, ktorého konkrétne hodnoty sa dozvedáme až po realizovaní náhodných vplyvov. Odchýlky skutočného dopytu po výrobkoch majú za následok dodatočné náklady na ich kompenzáciu. Tieto dodatočné náklady opäť znižujú vytvorený zisk. Je zrejme, že táto časť dodatočných nákladov môže byť čiastočne eliminovaná zapracovaním informácií o náhodnom dopyte do deterministickeho modelu.

Keď poznáme pravdepodobnostné rozdelenie dopytu, môžeme vytvárať špeciálne pravdepodobnostné ohraničenia (napr. môžeme požadovať, aby boli splnené v priemere alebo s vysokou pravdepodobnosťou), ktoré možno niekedy transformovať na deterministicke ekvivalent.

Dvojetapová formulácia spočíva v tom, že v druhej etape nachádzame taký spôsob kompenzácie vzniknutých odchýlok v dopyte, ktoré vedú najmenším kompenzačným nákladom. Je evidentné, že ak by bol predbežný výrobný program v prvej etape určený iný, boli by aj odchýlky v dopyte iné, a teda aj náklady na ich kompenzáciu by boli iné. To znamená, že úloha druhej etapy vlastne implicitne formuluje požiadavky na hľadanie optimálneho výrobného programu už v prvej etape.

Úlohou je nájsť taký predbežný výrobný program (pred realizáciou dopytu), ktorý zabezpečí za daných podmienok a znalostiach o dopyte minimum celkových nákladov za obe etapy, pričom náklady druhej etapy budeme uvažovať priemerné.

V literatúre DUPAČOVÁ, J., HURT, J. and ŠTĚPÁN, J., (2002) je preskúmaných množstvo variantov dvojetapových úloh stochastického programovania. Pre prípad, ktorý skúmame aj v tomto príspevku, ide o rovnomerné rozdelenie nezávislých zložiek vektora pravých strán. D. B. Judin 1981 pre tento prípad ukázal spôsob konštrukcie deterministického modelu a dokázal, že deterministický ekvivalent takejto dvojetapovej úlohy stochastického programovania je úlohou kvadratického programovania.

V našom prípade je úlohou vypočítať vektor množstiev výroby  $(x_1, x_2, \dots, x_n)$  a podiel financovania z úverových zdrojov  $y$  tak, aby celkové náklady za prvú aj druhú etapu boli minimálne:

*minimalizovať*

$$N = \sum_{j=1}^n c_j x_j + r_D y F_v + \min E[Q(\mathbf{x}, \mathbf{t}(\omega))]$$

a podmienok

$$\sum_{j=1}^n z_j x_j - r_D y F_v \geq Z$$

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n c_j x_j - y F_v \leq F_v$$

$$0 \leq y \leq 2$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n$$

kde symbol  $E$  je operátor strednej hodnoty a  $Q(\mathbf{x}, \mathbf{t}(\omega))$  je určené vzťahom (11).

### Použitá literatúra

1. JUDIN, D.B., 1981: Matematické meódy riadenia pri neúplnej informácii, Alfa, Bratislava
2. DUPAČOVÁ, J., HURT, J. and ŠTĚPÁN, J., 2002: Stochastic Modeling in Economics and Finance, Part II. Kluwer Acad. Publ., Dordrecht.
3. VLACHYNSKÝ, K. a kol., 2009: Podnikové financie. 1.vyd., EKONOMIA, Bratislava
4. BREALEY, R.A., MYERS, S.C., ALLEN, F., 2014: Teorie a praxe firemních financí. Aktualizované vydanie. BizBooks, Brno
5. HOZLÁR, E., 2016: Faktoring ako alternatívny zdroj financovania prevádzky podniku. Finančný manažér. - Roč. 16, č. 3 (2016), s. 30-42
6. Hozlár, E., 2017: Stochastický model štruktúry financovania podniku. In: Využitie kvantitatívnych metód vo vedecko-výskumnej činnosti a v praxi 12 : zborník zo seminára. - Bratislava : EKONÓM, S. 82-88

Tento príspevok vznikol v rámci riešenia projektu APVV-14-0020 Modelovanie Pareto optimálnych interakcií finančného a podnikového sektora v adaptačných procesoch a grantu

VEGA1/0340/16 Modelovanie efektívnosti financovania a kooperácie sektorov a regiónov slovenskej ekonomiky.

**Kontaktné údaje**

doc. RNDr. Eduard Hozlár, CSc.

Univerzita Komenského v Bratislave, FSEV, Ústav ekonómie

Mlynské luhy 4, 82105 Bratislava

Tel: (421 2) 2066862

email: [hozlar@fses.uniba.sk](mailto:hozlar@fses.uniba.sk)

# PRETRVÁVANIE REGIONÁLNYCH DISPARÍT NA SLOVENSKU

## PERSISTENCE OF THE REGIONAL DISPARITIES IN SLOVAKIA

*Michaela Chocholatá*

### **Abstrakt**

Predmetom príspevku je poukázanie na pretrvávajúce značné regionálne disparity v podmienkach jednotlivých okresov Slovenskej republiky v období 2001-2014 na báze nástrojov priestorovej analýzy dát. Predmetom analýzy je 5 socioekonomických indikátorov – priemerná nominálna mesačná mzda, Sauvyho index starnutia, miera nezamestnanosti, priemerné percento pracovnej neschopnosti a migračné saldo. Na báze nástrojov priestorovej exploratórnej analýzy ESDA (Exploratory Spatial Data Analysis) boli potvrdené značné regionálne rozdiely v hodnotách analyzovaných ukazovateľov ako aj ich pretrvávajúce v skúmanom období.

*Kľúčové slová: disparity, Slovensko, ESDA*

### **Abstract**

This paper deals with the persistence of the substantial regional disparities in individual districts of the Slovak Republic during the period 2001-2014 based on instruments of spatial data analysis. The paper analyses 5 socioeconomic indicators – average nominal monthly wage, Sauvy's index of aging, unemployment, average percentage of incapacity for work and net migration. Based on the tools of ESDA (Exploratory Spatial Data Analysis) were confirmed the significant regional differences in values of analysed indicators as well as their persistence during the analysed period.

*Keywords: disparity, Slovakia, ESDA*

## **1 INTRODUCTION**

The issue of regional disparities has become very popular and there exists huge amount of various studies dealing with this theme based on different approaches. In order to analyse regional disparities for differently defined regions various indicators can be used.

Especially attractive is the analysis of various socioeconomic indicators (GDP per capita, wages, unemployment, employment, education, crime etc.). Since these data have a spatial dimension, adequate has become to use the methods based on spatial data analysis. The issues of the regional economic growth, considering the spatial dependence can be found e.g. in studies of Arbia (2006), Paas et al. (2007), Furková and Chocholatá (2016), Chocholatá and Furková (2016). Furková (2016) analysed the spatial pattern of innovative activity in the EU regions based on exploratory spatial data analysis and spatial econometric approach. Chocholatá (2017) using the spatial data analysis dealt with the wage disparities across the Slovak districts. Hančlová and Tvrďý (2004) analysed the regional disparities across regions of the Czech Republic based on different socioeconomic indicators and different methods. Spurná (2008) provided the detailed characteristics of spatial autocorrelation and illustrated it for the three socioeconomic indicators in the regions of the Czech Republic. Matlovič and Matlovičová (2011) in their complex study dealt with the disparities in the Slovak Republic

based on different indicators and different methods, they also provide an extensive survey of geographical, economic and sociological studies dealing with the issue of disparities in Slovakia.

The main aim of the paper is to perform the spatial analysis of regional disparities for 79 Slovak districts based on the following socioeconomic indicators: average nominal monthly wage per employee, Sauvy's index of aging, unemployment, average percentage of incapacity for work and net migration.

The rest of the paper is organized as follows: Section 2 deals with methodology, Section 3 presents data and empirical results and Section 4 concludes.

## 2 METHODOLOGY

The regional disparities can be assessed based on different approaches. Since this paper concentrates on the instruments of spatial data analysis, we will briefly describe them in this section.

The ESDA (Exploratory Spatial Data Analysis) enables to investigate whether there exist some spatial pattern in the analysed data. As pointed out by Mitchell (2013), the graphical analysis based on plots should be in case of spatial data added by visualisation via maps (box maps, percentile maps, etc.). To gain information about the existence of spatial autocorrelation (= spatial dependence), i.e. correlation of a variable with itself through space, various tests on global and/or local level can be used. Global statistics (a single value for the entire data set) provide a measurement of the global spatial autocorrelation, local statistics usually assess the spatial autocorrelation for one particular region. The empirical part of this paper uses the global Moran's  $I$  statistics and the local Moran's  $I$  statistics (for both formulas see e.g. Fischer and Getis, 2010).

The important role in calculating the mentioned statistics plays the determination of nearby regions. The spatial relationships are formally expressed by the spatial weights matrix  $\mathbf{W}$  of dimension  $(n \times n)$ , where  $n$  is the number of regions in the data set and each element  $w_{ij}$  reflects the spatial influence of unit  $j$  on unit  $i$ . Regarding the different definitions of the spatial weight matrix  $\mathbf{W}$  – see e.g. Fischer and Getis, 2010. Calculations in this paper are based on a contiguity weight matrix of queen case definition of neighbours (two regions are neighbours in this sense if they share any part of a common border).

## 3 DATA AND EMPIRICAL RESULTS

The empirical part of the paper is based on the data for 79 Slovak districts (LAU1- Local Administrative Units) during the period 2001-2014 retrieved from the DATAcube database of the Statistical Office of the Slovak Republic (<http://datacube.statistics.sk/>). Analyses were performed in the free downloadable software GeoDa (<https://geodacenter.asu.edu/software/downloads>), MS Excel and EVIEWS. The shape file for the Slovak districts was downloaded from the following web-site: <http://wiki.freemap.sk/HraniceAdministrativnychUzemí>.

In order to analyse the regional disparities the following socioeconomic indicators were used: average nominal monthly wage per employee (in €), Sauvy's index of aging (the number of persons 65+ years old per 100 persons aged 0 to 14), unemployment (in %), average



percentage of incapacity for work (in %) and net migration (difference between immigrants and emigrants during the year to the average population in that year, per 1000 persons, i.e. in ‰).

The left part of the Fig. 1 presents the average values (calculated for the Slovak Republic as a whole) of individual indicators together with the maximum and minimum values (calculated for particular districts) during the period 2001-2014. All five indicators reflect substantial differences among the Slovak districts during the analysed period.

The average nominal wage in the Slovak Republic rose from 423.69 € in 2001 to 957.00 € in 2014. As it is clear from the Fig. 1, the regional wage disparities were huge – maximum values were reached for the district Bratislava II (with the exception of the year 2001), minimum values were recorded for the districts in the Eastern Slovakia. The biggest difference of 890.20 € between districts with maximum and minimum wage was detected in 2009 (Bratislava II: 1355.76 € and Bardejov: 465.56 €).

The second indicator, the Sauvy's index of aging, had during the analysed period in average a rising trend from 60.84 to 91.17 indicating a growing number of persons aged 65+ per 100 persons aged 0 to 14. The districts with the smallest proportion of elderly people per 100 children were Košice III, Námestovo and Kežmarok, while the highest proportion of elderly people per 100 children lived in the district Bratislava I.

Enormous disparities are also observable for unemployment rate. Since the Slovak average values started at 18.63 % in 2001 decreasing to 7.99 % in 2007 and reaching the value of 12.29 % in 2014, the minimum values (various districts of Bratislava) were far below the average and the maximum values reached for the district Rimavská Sobota highly exceeded the average level.

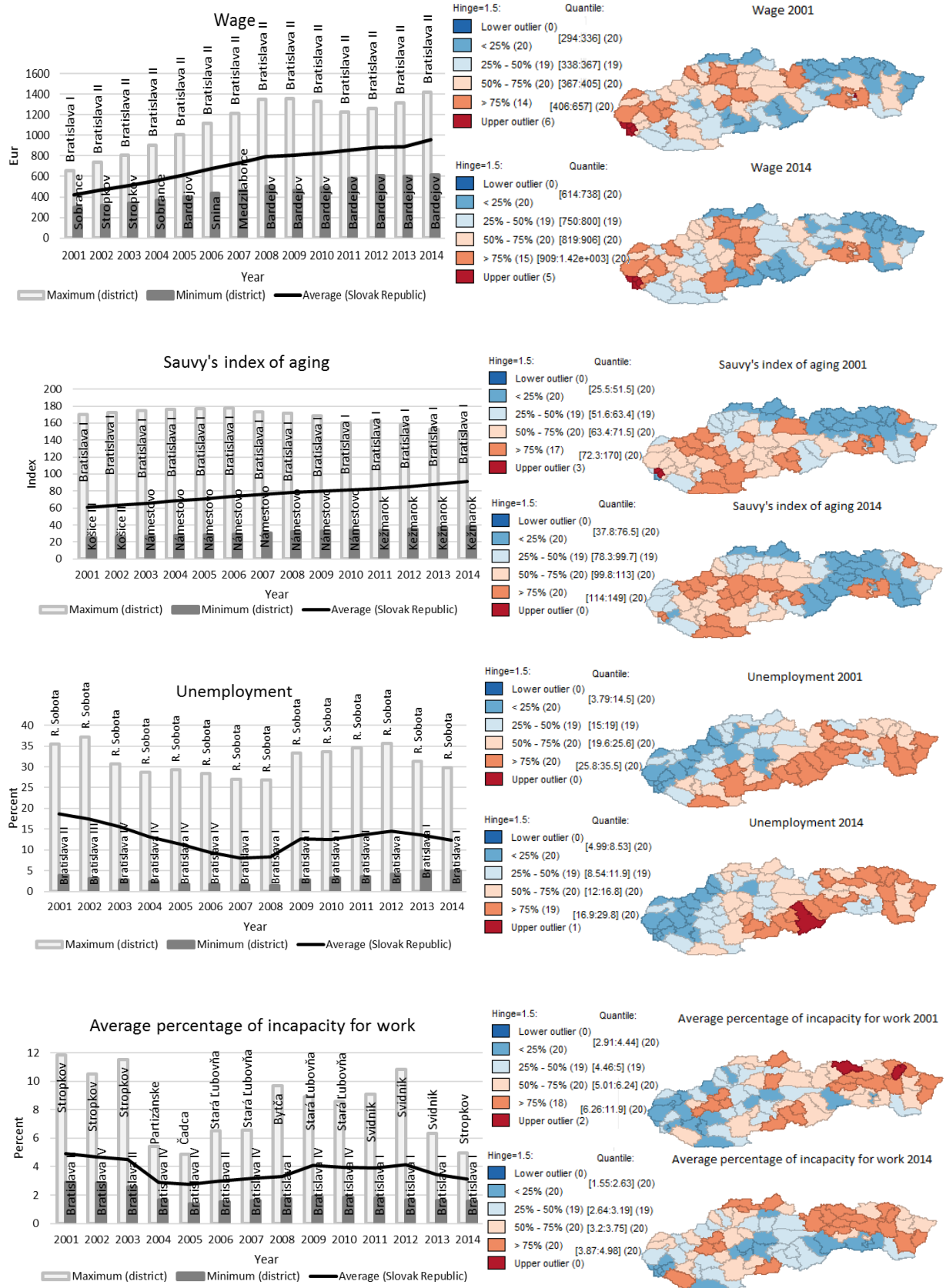
The next indicator reflects the average percentage of incapacity for work starting with 4.905 % in 2001 without clearly up- or downward trend ending with 3.107 % in 2014. Since the lowest percentage was observed for the various districts of Bratislava, the highest values were reached (with three exceptions) for the districts situated in the Eastern Slovakia.

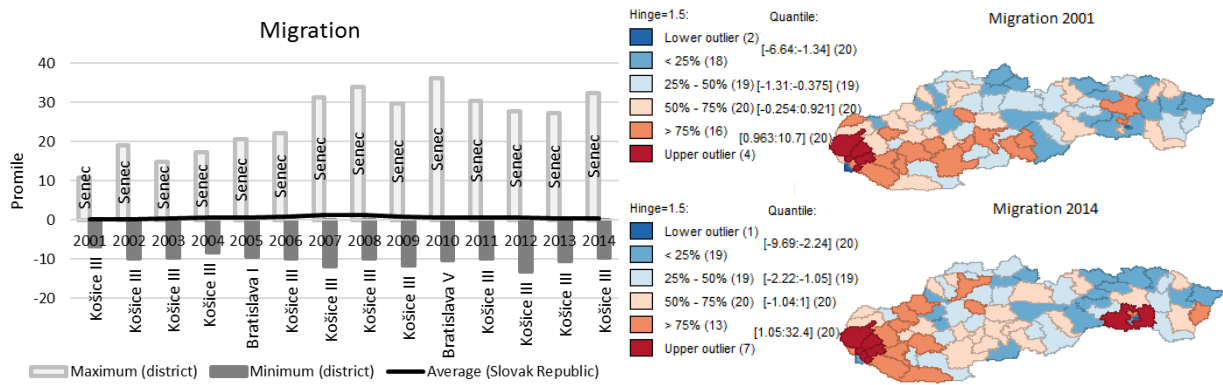
Concerning the indicator of migration, the average values for the Slovak Republic were slightly above 0 during the whole analysed period. The highest positive values were reached for district Senec, while the lowest negative values were recorded mostly for the district Košice III.

The right part of the Fig. 1 based on box maps constructed for the years 2001 and 2014 enables to gain further insight into the development of analysed indicators across individual districts.

The districts with high average wages are mainly the districts with low level of unemployment rate and vice versa. The districts with the smallest proportion of elderly people per 100 children are situated mainly in the northern, south-eastern and eastern part of Slovakia. Furthermore, it is possible to observe that the average percentage of incapacity for work is high especially in the districts with low wages and high unemployment. Fig. 1 also confirms the commonly known fact that the highest is the net migration into the districts situated in the western part of Slovakia with very low level of unemployment.

Fig. 1: Visualisation of selected indicators' data based on average, maximum and minimum values for 2001-2014 (left part) and box maps for 2001 and 2014 (right part)

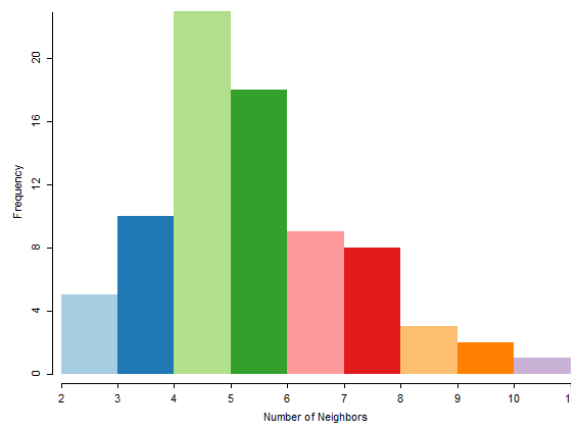




Source: Author's illustration

Concerning the spatial aspect of analysis, the empirical part begins with the ESDA in order to examine the fact that the values of analysed indicators in one district may be associated with their values in neighbouring districts and to confirm the existence of spatial clustering. The corresponding spatial weight matrix  $\mathbf{W}$  was specified as a contiguity weight matrix of queen case definition of neighbours (for the connectivity histogram indicating the number of neighbours see Fig. 2).

Fig. 2: Number of neighbours based on queen contiguity weight matrix



Source: Author's illustration

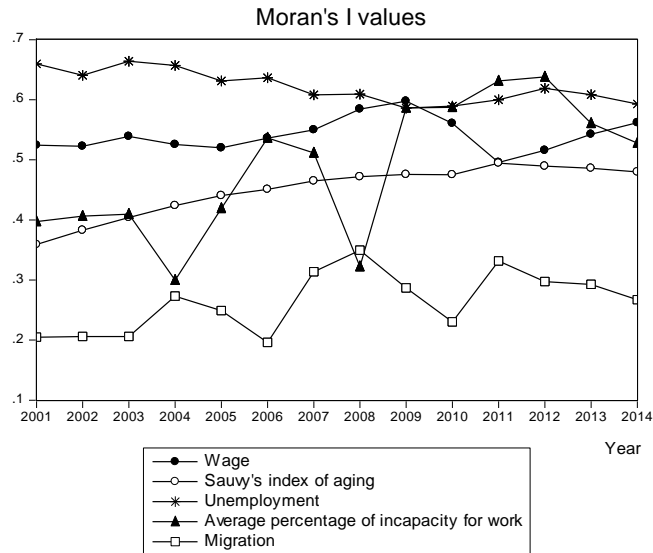
To assess the existence of spatial autocorrelation the global Moran's  $I$  statistics were calculated. The development of global Moran's  $I$  statistics for individual analysed indicators across the period 2001-2014 is graphically illustrated in Fig. 3.

Since the values of global Moran's  $I$  statistics were in all analysed cases larger than the expected value  $E(I) = -1/(n-1) = -0.01282$ , the presence of the positive spatial autocorrelation was confirmed<sup>1</sup>, i.e. it is much more likely that the districts with high (low) value of the concrete indicator will have neighbours with also high (low) value of this

<sup>1</sup> Randomization approach taking into account 999 permutations was used to confirm the statistical significance of results.

indicator than in case of pure randomness. In this context, it is further necessary to point out that the global Moran's  $I$  only provides a measurement of global spatial association. In the next step the local spatial analysis was provided based on calculation of the local Moran's  $I$  statistic.

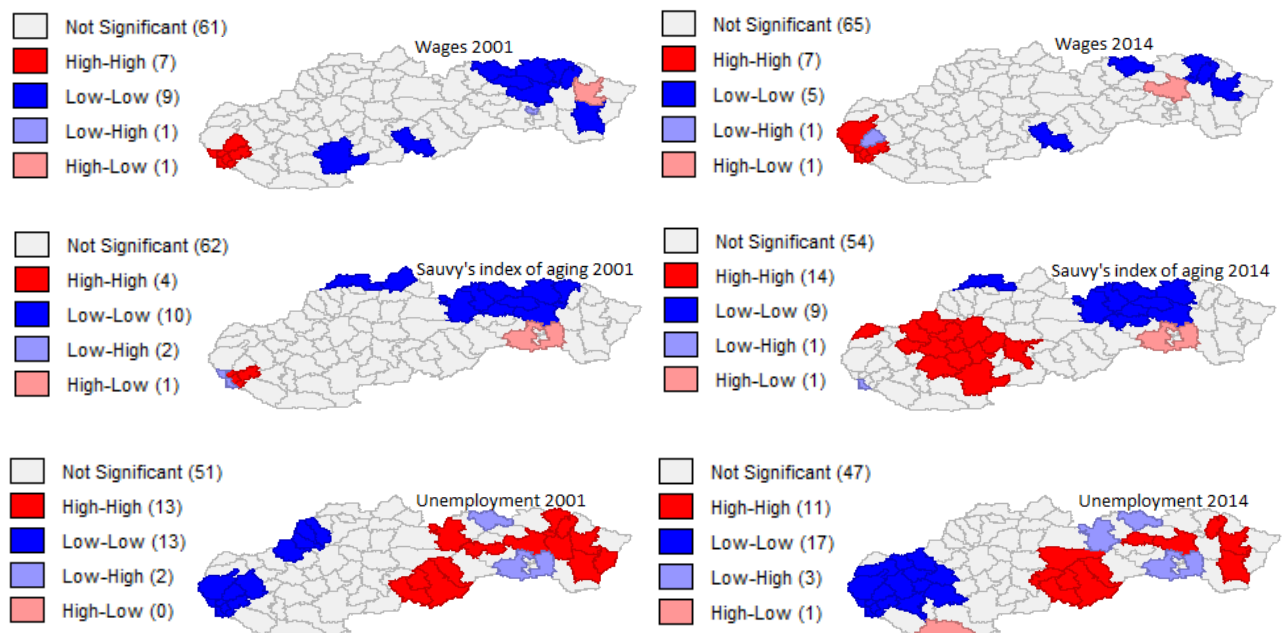
Fig. 3: Development of Moran's  $I$  values for analysed indicators across the period 2001-2014

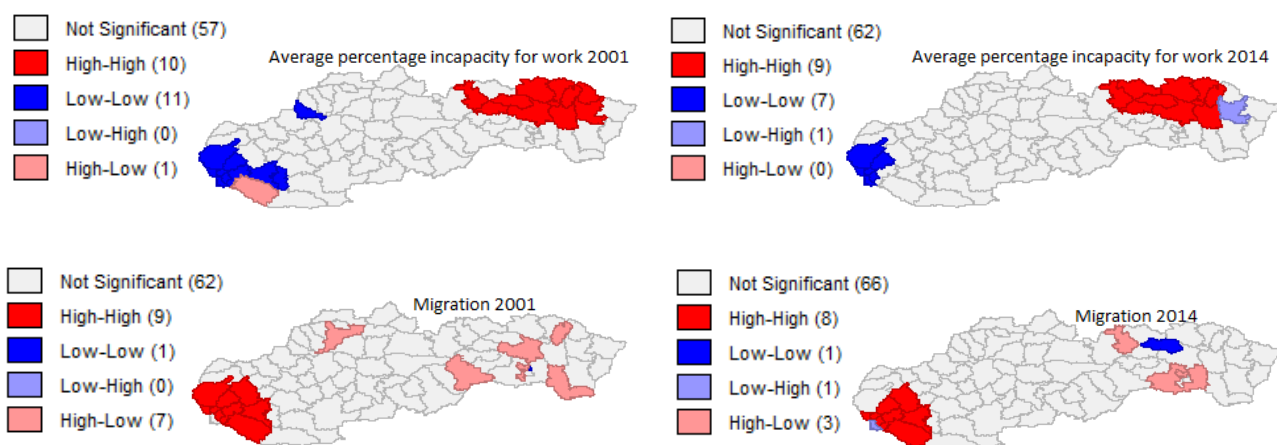


Source: Author's illustration

The software GeoDa enables to map the LISA (Local Indicators of Spatial Association) clusters which show the districts with significant spatial association (for the 2001 and 2014) – see Fig. 4. It can be distinguished districts with the positive spatial autocorrelation of the type high-high and low-low as well as districts with the negative autocorrelation of the type high-low and low-high.

Fig. 4: LISA cluster maps for the years 2001 and 2014





Wages – Fig. 4 shows particular high values in districts of Bratislava and surrounding districts and low values mostly in the regions located in the eastern part of Slovakia. Sauvy’s index of aging – since the situation with statistically significant high-high clusters changed over time dramatically, the low-low clusters remained almost the same. Unemployment – with high-high clusters prevailing in the western part of Slovakia and low-low clusters located mostly in the eastern and middle southern part of Slovakia. Average percentage of incapacity for work – high values are found in the northern districts of East Slovakia and low values prevailed in the districts located in the western part of Slovakia. Migration – high levels were confirmed to be significant in the western part of the country.

## 4 CONCLUSION

The analysis of the regional disparities across districts of the Slovak Republic during the period 2001-2014 based on instruments of spatial data analysis has proved significant differences among the individual districts. The values of all the 5 socioeconomic indicators – average nominal monthly wage, Sauvy's index of aging, unemployment, average percentage of incapacity for work and net migration showed significant positive spatial autocorrelation, i.e. existence of the spatial spillover effects. Based on the LISA cluster maps it was possible to further identify the statistically significant clusters. Since the significant regional differences in values of analysed indicators as well as their persistence during the analysed period was proved, the issue of disparities diminishing is still being a highly actual issue in the Slovak Republic.

### References

1. ARBIA, G. 2006. Spatial econometrics: statistical foundations and applications to regional convergence: with 19 figures and 11 tables. Berlin: Springer.
2. CHOCHOLATÁ, M., and FURKOVÁ, A. 2016. Does the location and institutional background matter in convergence modelling of the EU regions? Central European Journal of Operations Research. Retrieved September 05, 2016, from <http://link.springer.com/article/10.1007/s10100-016-0447-6?view=classic> .
3. CHOCHOLATÁ, M. 2017. Analýza mzdových disparít v okresoch Slovenska v období 2001-2015. In AIESA – budovanie spoločnosti založenej na vedomostiach. Bratislava, Slovakia: Ekonóm, pp. 316-326.
4. DATAcube. (n.d.). Retrieved October 09, 2017, from <http://datacube.statistics.sk/>

5. FISCHER, M.M., and GETIS, A. 2010. Handbook of Applied Spatial Analysis. Software Tools, Methods and Applications. Springer-Verlag, Berlin, Heidelberg, 2010.
6. FURKOVÁ, A. 2016. Spatial pattern of innovative activity in the EU regions : exploratory spatial data analysis and spatial econometric approach. In Advances in applied business research : the L.A.B.S. initiative. (pp. 159-186). New York: Nova Science.
7. FURKOVÁ, A., and CHOCHOLATÁ, M. 2016. Spatial econometric modelling of regional club convergence in the European Union. In Ekonomický časopis, vol. 64, No. 4, pp. 367-386.
8. HANČLOVÁ, J., and TVRDÝ, L. 2004. Classification of the Regions. In Multiregional and Regional Models. Faculty of Economics, Technical University of Ostrava.
9. MATLOVIČ, R., and MATLOVIČOVÁ, K. 2011. Regionálne disparity a ich riešenie na Slovensku v rozličných kontextoch. In Folia Geographica 18, Prešov, pp. 8-87.
10. MITCHELL, W. 2013. Introduction to spatial econometric modelling. In Centre of Full Employment and Equity Working Paper, No. 1–13.
11. (n.d.). Retrieved October 9, 2017, from <http://wiki.freemap.sk/HraniceAdministrativnychUzemi>
12. (n.d.). Retrieved October 09, 2017, from <https://geodacenter.asu.edu/software/downloads>
13. PAAS, T. et al. 2007. Econometric Analysis of income convergence in selected EU countries and their NUTS 3 level regions. Retrieved February 05, 2015, from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1078863](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1078863)
14. SPURNÁ, P. 2008. Prostorová autokorelace – všudypřítomný jev při analýze prostorových dat? In Sociologický časopis/Czech Sociological Review, 2008, vol. 44, No. 4, pp. 767-787.

### **Acknowledgements**

This work was supported by the Grant Agency of Slovak Republic – VEGA grant No. 1/0248/17 "Analysis of Regional Disparities in the EU based on Spatial Econometric Approaches".

### **Contact address**

doc. Ing. Michaela Chocholatá, PhD.  
 Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky  
 Dolnozemska 1/b, 852 35 Bratislava  
 Tel: (421 2) 67 295 832  
 Email: [michaela.chocholata@euba.sk](mailto:michaela.chocholata@euba.sk)

# STRATEGIE „NEJBLIŽŠÍ BERE“ V UMISŤOVACÍCH ÚLOHÁCH

## STRATEGY “THE NEAREST ONE WINS” IN LOCATION PROBLEMS

*Jaroslav Janáček*

### **Abstrakt**

Strategie „nejbližší bere“ je v umisťovacích úlohách s účelovou funkcí  $p$ -mediánového typu zabezpečena samotným minimalizačním procesem. V tomto příspěvku se zabýváme o mnoho těžším případem, kdy je maximalizována  $p$ -mediánová účelová funkce vhodnou volbou umístění středisek obsluhy a přitom má být uspokojovaný požadavek přiřazen nejbližšímu umístěnému středisku. Navrhujeme zde konstrukci podmínek pro zabezpečení uvedené strategie a studujeme vliv těchto podmínek na výpočetní dobu potřebnou pro řešení daného typu úloh.

***KLíčové slová:** strategie „nejbližší bere“, umisťovací úloha, střediska obsluhy*

### **Abstract**

The strategy “the nearest one wins” is usually ensured by the very minimization process, which solves  $p$ -median like location problem. Within this paper, we deal with a much more difficult case, when the  $p$ -median objective function is maximized by service center location submit to the rule that a service demand must be assigned to the nearest located service center. We suggest a way of constraint construction to ensure the above-mentioned strategy and to study an impact of the constraint on the computational time necessary for solution of such problem.

***Keywords:** strategy “the nearest one wins”, location problem, service centers*

## **1 ÚVOD**

Klasický návrh záchranného obslužného systému spočívá v umístění daného počtu obslužných středisek v některých z možných umístění tak, aby bylo minimalizováno zatížení průměrného uživatele v obsluhovaném regionu. O zatížení uživatele je předpokládáno, že je úměrné jeho vzdálenosti od nejbližšího obslužného střediska. Prvotní návrh umístění obslužných středisek bývá vypracován administrátorem systému, zastupujícím veřejnost, tedy uživatelé systému [1], [4]. Vlastní poskytování služby bývá na základě výsledků konkurzů svěřeno soukromým poskytovatelům dané služby. Tak v daném regionu, jeden poskytovatel vlastní několik středisek a jeho záchranná vozidla jsou přidělována náhodně vzniklým požadavkům uživatelů administrátorem podle pravidla, že požadavek na poskytnutí služby je přiřazen nejbližšímu středisku. Samotný příjem poskytovatele služby je potom úměrný dopravní práci potřebné pro uspokojení daného požadavku, tedy jinými slovy, příjem poskytovatele je úměrný vzdálenosti, kterou jeho vozidla projedou při uspokojování jemu přiřazených požadavků na obsluhu.

Protože se každý region v čase vyvíjí jak z hlediska rozložení obyvatel, tak i z hlediska frekvence požadavků, musí být systém po nějakém čase revidován. Podle typu systému a legislativy státu, v kterém systém zabezpečuje danou službu, je „reengineering“, tedy změna poloh středisek, navržena znovu administrátorem systému [3], [6] a [7] anebo je navrhována jednotlivými soukromými poskytovateli [5], ovšem za podmínek stanovených

administrátorem. Administrátor má vhodným nastavením podmínek chránit zájmy veřejnosti (např. co nejkratší průměrnou dobu dojezdu na místo nehody), které jsou ale v konfliktu s ekonomickým zájmem poskytovatele.

V tomto příspěvku se budeme zabývat modelováním chování poskytovatele služby při „reengineeringu“ jeho části obslužného systému. Vyřešíme zde na úrovni modelu problém, který vznikne při formulaci úlohy maximalizace příjmu poskytovatele služby prostřednictvím přemístování středisek, kdy je třeba zabezpečit, aby požadavky střediskům byly přiřazovány podle pravidla „nejbližší bere“. Prostřednictvím numerických experimentů zjistíme, jak se projeví použití příslušných podmínek na zvýšení doby výpočtu optimálního řešení dané úlohy oproti době řešení umísťovací úlohy, která nevyžaduje speciální podmínky pro zabezpečení pravidla „nejbližší bere“.

## 2 AKTUALIZACE ZÁCHRANNÉHO SYSTÉMU PODLE ADMINISTRÁTORA

Pro formulaci úlohy systémově optimální aktualizace záchranného systému použijeme následujícího označení. Symbolem  $J$  označíme množinu potenciálních uživatelů (např. obcí) systému, kde  $b_j$  označuje četnost požadavků na obsluhu pro  $j \in J$ . Symbol  $I$  označuje množinu všech možných umístění středisek v obsluhovaném regionu. Symbol  $d_{ij}$  je celočíselná vzdálenost míst  $i$  a  $j$ , kde  $i, j \in I \cup J$ . Maximální relevantní vzdálenost je  $m$ . Současná struktura záchranného systému je určena dvěma disjunktními množinami středisek  $I_L \subset I$  a  $I_F \subset I$ , kde  $I_L$  obsahuje  $p$  středisek poskytovatele, které může administrátor přemístit a kde množina  $I_F$  je množina středisek ostatních poskytovatelů, kteří se na aktualizaci nepodílí.

Administrátor systému dbá omezení maximální vzdálenosti  $D$  mezi současným umístěním střediska a mezi možným novým umístěním. Abychom byli schopni formulovat model, zavedeme množinu středisek  $N_t = \{i \in I - I_F : d_{it} \leq D\}$ , do kterých může být středisko  $t \in I_L$  přesunuto. Podobně množina  $S_i = \{t \in I_L : i \in N_t\}$  obsahuje všechna střediska uvažovaného poskytovatele, která mohou být přesunuta do daného  $i \in I_R$ . Podmnožina  $I_R \subseteq I - I_F$  je definovaná vztahem  $I_R = \bigcup_{t \in I_L} N_t$ . Dále zavedeme koeficienty  $a^s_{ij}$  pro každou dvojici  $i, j \in I_R \cup I_F$  a  $j \in J$ , kde  $a^s_{ij} = 1$  právě tehdy, když  $d_{ij} \leq s$  a jinak  $a^s_{ij} = 0$  pro  $s = 0, 1, \dots, m-1$ .

Dále zavedeme rozhodovací proměnné. Proměnné  $y_i$  jsou definované pro každé  $i \in I_R$ . Proměnná  $y_i$  nabyde hodnoty 1, právě tehdy, když v možném umístění  $i$  má být umístěno středisko. Jinak proměnná nabývá hodnoty 0. Proměnná  $u_{ti} \in \{0, 1\}$  pro  $t \in I_L$  a  $i \in N_t$  nabyde hodnoty 1, pokud má být středisko přesunuto z  $t$  do  $i$ , jinak tato proměnná nabývá hodnotu 0. Nakonec zavedeme pomocné proměnné  $x_{js}$  pro  $j \in J$  a  $s = 0, 1, \dots, m-1$ , kde  $x_{js} = 1$  pokud v okruhu s poloměrem  $s$  okolo uživatele  $j$  není umístěno žádné středisko. Potom můžeme úlohu formulovat následujícím způsobem:

$$\text{Maximalizujte } \sum_{j \in J} b_j \sum_{s=0}^{m-1} x_{js} \quad (1)$$

$$\text{Za podmíněk } \sum_{i \in I_R} y_i = p \quad (2)$$

$$\sum_{i \in N_t} u_{ti} = 1 \text{ pro } t \in I_L \quad (3)$$

$$\sum_{t \in S_i} u_{ti} \leq y_i \text{ pro } i \in I_R \quad (4)$$



$$x_{js} + \sum_{i \in I_R} a_{ij}^s y_i + \sum_{i \in I_F} a_{ij}^s \geq 1 \quad \text{pro } j \in J, s = 0, \dots, m-1 \quad (5)$$

$$y_i \in \{0, 1\} \quad \text{pro } i \in I_R \quad (6)$$

$$u_{ti} \in \{0, 1\} \quad \text{pro } t \in I_L, i \in N_t \quad (7)$$

$$x_{js} \in \{0, 1\} \quad \text{pro } j \in J, s = 0, \dots, m-1 \quad (8)$$

Účelová funkce (1) vyjadřuje délku cest realizovanou vozidly záchranného systému. Podmínka (2) zabezpečuje, že počet středisek poskytovatele služby nebude změněn. Podmínky (3) dovolují přesunout středisko z pozice  $t$  do nejvýše jednoho jiného umístění v okruhu  $D$ . Podmínky (4) umožňují do  $i$  přenést nejvýše jedno středisko a to ze vzdálenosti nejvýše  $D$ . Podmínky (5) zabezpečují vztah mezi proměnnými  $y_i$  a  $x_{js}$  [2], [5].

### 3 AKTUALIZACE ZÁCHRANNÉHO SYSTÉMU PODLE POSKYTOVATELE SLUŽBY

Pro formulaci úlohy optimalizace příjmu poskytovatele služby za podmínek stanovených administrátorem veřejného záchranného systému použijeme stejná označení, proměnné i některé podmínky uvedené v předchozí sekci. Navíc zavedeme následující konstanty a proměnné. Nejdříve zavedeme koeficienty, jejichž hodnota bude úměrná příjmu poskytovatele služby za obsluhu požadavku uživatele  $j$ . Pro  $i \in I_R \cup I_F$  a  $j \in J$  bude  $c_{ij} = 0$  if  $d_{ij} \geq \min\{d_{ij}: t \in I_F\}$  a jinak  $c_{ij} = b_j d_{ij}$ . Spolu s koeficienty zavedeme přiřazovací proměnné  $z_{ij} \in \{0, 1\}$  pro  $i \in I_R \cup I_F$  a  $j \in J$ , kde  $z_{ij} = 1$  jestli uživatel  $j$  je přiřazen středisku  $i$ . Nakonec zavedeme dvě pomocné struktury označené jako  $\{P_j\}$  a  $\{R_j\}$ , kde  $j \in J$ . První struktura je soustava uspořádaných seznamů, kde seznam  $P_j$  umístění středisek  $i \in I_R$  je uspořádán tak, že platí následující nerovnosti  $d_{P_j(1)j} \leq d_{P_j(2)j} \leq \dots \leq d_{P_j(|I_R|)j}$ . Prvkem  $R_j$  z druhé struktury je uspořádaný seznam indexů z oboru  $1, \dots, |I_R|$ , kde  $R_j(r)$  je nejmenší index, pro který platí  $d_{P_j(r)j} < d_{P_j(R_j(r))j}$ . Zřejmě  $r+1 \leq R_j(r)$ . Potom můžeme novou úlohu formulovat následujícím způsobem:

$$\text{Maximalizujte } \sum_{j \in J} \sum_{i \in I_R} c_{ij} z_{ij} \quad (9)$$

Za podmínek (2) – (8)

$$\sum_{j \in J} b_j \sum_{s=0}^{m-1} x_{js} \leq F \quad (10)$$

$$\sum_{i \in I_R \cup I_F} z_{ij} = 1 \quad \text{pro } j \in J \quad (11)$$

$$z_{ij} \leq y_i \quad \text{pro } j \in J, i \in I_R \quad (12)$$

$$1 - y_{P_j(k)} \geq \sum_{r=R_j(k)}^{|I_R|} z_{P_j(r)j} \quad \text{pro } j \in J, k = 1, \dots, |I_R| - 1 \quad (13)$$

$$z_{ij} \in \{0, 1\} \quad \text{pro } i \in I_R \cup I_F, j \in J \quad (14)$$

Účelová funkce (9) vyjadřuje příjem uvažovaného poskytovatele služby. Pokud je uživatel blíže ke středisku jiných poskytovatelů, potom je příjem uvažovaného poskytovatele nulový. Přiřazení uživatele středisku poskytovatele, které je dále než nejbližší středisko brání podmínky (13). Podmínka (10) brání tomu, aby celková disutilita všech uživatelů přesáhla administrátorem stanovenou hodnotu  $F$ . Podmínky (11) zabezpečují, aby požadavek každého uživatele byl přiřazen právě jednomu středisku. Vazební podmínky (12) zabezpečují, že pokud bude požadavek uživatele  $j$  přiřazen možnému umístění střediska, tak tam bude

středisko vybudované. Za povšimnutí stojí podmínky (13), které brání optimalizačnímu procesu přiřadit požadavek uživatele  $j$  jinému středisku než je to nejbližší. Vliv dodání těchto podmínek k modelu na výpočetní náročnost získání optimální úlohy je studován v další sekci.

#### 4 NUMERICKÉ EXPERIMENTY

Dále uvedené numerické experimenty byly provedeny pomocí optimalizačního softwaru FICO Xpress 8.0 (64-bit, release 2016) na PC vybaveném procesorem Intel® Core™ i7 5500U s parametry: 2.4 GHz a 16 GB RAM. Použité úlohy byly odvozeny ze skutečného záchranného systému Slovenské republiky, která je rozčleněna na osm samosprávných krajů, tj. Bratislava (BA), Banská Bystrica (BB), Košice (KE), Nitra (NR), Prešov (PO), Trenčín (TN), Trnava (TT) a Žilina (ZA). Všechna obývaná místa byla brána v úvahu jako uživatelé a počet obyvatel byl považován za velikost požadavku  $b_j$ . Velikosti množin  $J$  a  $I$  byly v každém samosprávném kraji stejné a jejich hodnoty jsou uvedeny v sloupci „ $|I|$ “ tabulky 1. Celkový počet umístěných středisek je uveden v tabulce v sloupci „ $|I_I|$ “. Množina středisek uvažovaného poskytovatele byla generována náhodně v každém z deseti případů pro každý samosprávný kraj tak, aby jejich počet představoval jisté procento z celkového počtu středisek (viz sloupec „ $|I_I|$  [%]“ tabulky 1.). Hodnota účelové funkce (1) pro současný stav je uvedena v sloupci „Účelová funkce (1)“. V pravé části tabulky 1 jsou uvedeny údaje o optimalizaci úlohy (1)-(8). Zde sloupec „Doba [s]“ udává průměrnou dobu řešení v sekundách a sloupec „Dec. [%]“ pokles účelové funkce v procentech oproti současnému stavu.

Tabulka 1: Současný stav a výsledky optimalizace úlohy (1)-(8)

Kraj			Současný stav		Optimalizovaný stav	
	$ I $	$ I_I $	Účelová funkce (1)	$ I_I $ [%]	Doba [s]	Dec. [%]
BA	87	14	21842	55.1	0.02	6.34
BB	515	36	32476	44.9	0.21	2.40
KE	460	32	36363	46.0	0.36	3.21
NR	350	27	38831	50.7	0.48	3.94
PO	664	32	42740	44.3	0.28	1.59
TN	276	21	26683	52.9	0.12	2.50
TT	249	18	31582	49.6	0.13	4.92
ZA	315	29	31955	46.8	0.11	3.49

Následující experimenty byly zaměřeny na zjištění toho, jak dodání podmínek (13) k modelu ovlivní výpočetní náročnost získání optimální úlohy.

Tabulka 2: Průměrné výpočetní časy optimálního řešení úlohy optimalizace příjmu poskytovatele pro různě nastavené hodnoty  $F$  v procentech z absolutní hodnoty poklesu účelové funkce úlohy (1)-(8)

Kraj\Redukce $F$ o procenta	0	20	40	60	80	100
BA	3.5	3.0	2.5	2.0	1.5	0.2
BB	76.1	52.5	40.4	28.2	14.0	6.3
KE	133.0	106.8	62.8	46.1	18.6	10.5
NR	109.9	65.5	49.5	29.9	17.7	5.6
PO	88.7	51.6	40.5	42.0	19.5	11.3
TN	24.1	21.6	13.5	9.9	6.0	2.4
TT	142.6	95.6	52.7	25.9	13.0	2.3
ZA	99.7	69.4	56.8	28.7	11.6	2.9
<b>Průměr</b>	<b>84.7</b>	<b>58.2</b>	<b>39.8</b>	<b>26.6</b>	<b>12.8</b>	<b>5.2</b>

Zde byl každý případ řešen pro šest různých nastavení hodnoty  $F$  na hodnotu, kdy účelová funkce zjištěná pro současný stav byla redukována o jisté procento největšího možného snížení vypočítaného řešením úlohy (1)-(8). Procenta snížení byla volena takto: 0, 20, 40, 60, 80 a 100.

## 5 ZÁVĚR

V příspěvku jsme ukázali, jak lze v modelu umisťovací úlohy maximalizující příjem poskytovatele služby zabezpečit sérií speciálních podmínek to, aby příjem nebyl dosahován na základě přiřazení požadavku střediskám, které jsou umístěna od požadavku dále než nejbližší možnost jeho uspokojení. Prověřili jsme i praktickou možnost použití navržených podmínek při řešení úloh reálného rozsahu a to provedením numerických experimentů. Experimenty ukázaly, že i když podmínky zvyšují nárok na dobu výpočtu optimálního řešení o několik řádů, tak doba výpočtu nepřesahující tři minuty je stále akceptovatelná.

## OZNÁMENÍ

Tato práce vznikla v rámci projektů VEGA 1/0518/15 “Spolahlivé záchranné systémy s neistou dosiahnuteľnosťou služby“, VEGA 1/0463/16 “Ekonomicky efektívna prevádzka elektrických vozidiel v inteligentných mestách a komunitách“ a APVV-15-0179 “Spolahlivosť záchranných systémov na infraštruktúre s neistou funkcionalitou kritických prvkov”.

### Použitá literatúra

1. BROTCORNE, L., LAPORTE, G., SEMET, F. 2003. Ambulance location and relocation models. *European Journal of Operational Research* 147, pp. 451–463
2. GARCÍA, S., LABBÉ, M., MARÍN, A. 2011. Solving large  $p$ -median problems with a radius formulation. *INFORMS Journal on Computing*, Vol. 23, No 4, pp. 546-556
3. GUERRIERO, F., MIGLIONICO, G., OLIVITO, F. 2016. Location and reorganization problems: The Calabrian health care system case. *European Journal of Operational Research* 250, pp. 939-954
4. JÁNOŠÍKOVÁ, L., ŽARNAY, M. 2014. Location of emergency stations as the capacitated  $p$ -median problem. In *International scientific conference: Quantitative Methods in Economics-Multiple Criteria Decision Making XVII*, Virt, Slovakia
5. KVET, M., JANÁČEK, J. 2016. Radiálny prístup na zlepšenie existujúceho záchranného systému. In *Optimalizační úlohy v dopravních a logistických systémech a SW podpora rozhodování v inteligentních dopravních systémech*, Praha, Czech Republic, pp. 11-25
6. REUTER-OPPERMANN, M., VAN DEN BERG, P. L., VILE, J. L. 2017. Logistics for Emergency Medical Service systems. *Health Systems*, Vol. 6, No 3, pp 187-208
7. SAYAH, D., IRNICH, S. 2016. A new compact formulation for the discrete  $p$ -dispersion problem. *European Journal of Operational Research*, Vol. 256, No 1, pp. 62-67

### Kontaktné údaje

prof. RNDr. Jaroslav Janáček, CSc.

Žilinská univerzita v Žiline, Fakulta riadenia a informatiky

Univerzitná 1, 010 26 Žilina

Tel: +421 (0)41 513 4204

email: [jaroslav.janacek@fri.uniza.sk](mailto:jaroslav.janacek@fri.uniza.sk)

# ANALÝZA ZLOŽIEK UKAZOVATEĽA MEDZINÁRODNEJ INVESTIČNEJ POZÍCIE A ICH VPLYV NA HRUBÝ DOMÁCI PRODUKT

## THE ANALYSIS OF THE INTERNATIONAL INVESTMENT POSITION INDICATOR'S COMPONENTS AND THEIR EFFECT ON GROSS DOMESTIC PRODUCT

*Miroslava Jánošová*

### **Abstrakt**

Predkladaný príspevok sa venuje analýze jednotlivých zložiek ukazovateľa medzinárodnej investičnej pozície a ich vplyvu na ekonomický rast Slovenska, vyjadrený ako zmena hrubého domáceho produktu. Tieto vplyvy sú skúmané prostredníctvom ekonometrického modelu založenom na princípe ECM (Error correction model). Článok je členený na niekoľko samostatných častí, v prvej časti sa venujeme popisu jednotlivých zložiek ukazovateľa medzinárodnej investičnej pozície. V druhej časti sa zameriame na popis modelu ECM. Tretia posledná časť je venovaná výsledkom odhadnutého modelu.

***Kľúčové slová:** medzinárodná investičná pozícia, model s korekčným členom, hrubý domáci produkt*

### **Abstract**

Presented paper deals with the analysis of the individual components of the international investment position indicator and of their impact on economic growth of Slovakia, manifested as the change of gross domestic product. These effects will be examined through econometric model based on the methodology of ECM (Error correction model). The paper is structure into several separate sections. First section is dedicated to the description of the individual components of international investment position indicator. In the second section we focus on the description of the ECM model. The third and final section presents the results of the estimated model.

***Keywords:** international investment position, error correction model, gross domestic product*

## **1 ÚVOD**

Ekonomický systém je založený na výmene tovarov a služieb a na expanzii finančného kapitálu. Cieľom predkladaného príspevku je skúmanie vplyvu ukazovateľov, ktoré tvoria relatívne nový ukazovateľ MIP, na hrubý domáci produkt Slovenska. Samotný ukazovateľ MIP spolu s platobnou bilanciou zobrazuje vzťah medzi rezidentmi a nerezidentmi krajiny. Tento ukazovateľ je svojím spôsobom formu bilancie, ktorej úlohou je porovnanie stavu investícií rezidentov domácej ekonomiky v zahraničí, so stavom investícií zahraničných subjektov v domácej ekonomike. Skúmanie ukazovateľa MIP z hľadiska jeho vplyvu na ekonomický rast krajiny meraného hrubým domácim produktom považujeme za dôležitý z dôvodu trendu liberalizácie kapitálových tokov a potreby hľadania vhodného a významného finančného indikátora.

## **2 CHARAKTERISTIKA ZLOŽIEK UKAZOVATEĽA MEDZINÁRODNEJ INVESTIČNEJ POZÍCIE**

Ukazovateľ MIP predstavuje doplnkovú štatistiku platobnej bilancie a z tohto dôvodu sú niektoré jeho základné komponenty zhodné s finančným účtom platobnej bilancie. MIP sa primárne člení na aktíva (investície smerujúce z krajiny), pasíva (investície smerujúce do krajiny) a na saldo, ktoré predstavuje rozdiel medzi aktívami a pasívami, a teda vyjadruje čistú investičnú pozíciu. V predkladanom príspevku sa zameriame len na pasíva zložiek ukazovateľa MIP, a teda len na tie investície, ktoré smerujú do krajiny. V nasledujúcej časti si jednotlivé zložky ukazovateľa MIP bližšie charakterizujeme. MIP je do svojej agregovanej podoby napočítaný z ukazovateľov priamych zahraničných investícií (PZI), portfóliových investícií (PORTI), finančných derivátov (FD) a ostatných investícií (OSTI)<sup>1</sup>.

### ***Priame zahraničné investície***

Ukazovateľ priame zahraničné investície tvoria kategóriu medzinárodných investícií, ktoré vyjadrujú zámer subjektu, ktorý je rezidentom jednej ekonomiky, získať trvalý podiel v podniku so sídlom v inej ekonomike. Pri určovaní platobnej bilancie/medzinárodnej investičnej pozície eurozóny sa v súlade s medzinárodnými štandardmi (MMF) pri určovaní existencie vzťahu priamej investície používa "kritérium 10 % vlastníckeho podielu", t. j. trvalého podielu.

### ***Portfóliové investície***

Portfóliové investície v rámci medzinárodnej investičnej pozície sa určujú výlučne na základe koncoročných údajov o pozíciách v investičných cenných papieroch, krátko a dlhodobých cenných papierov emitovaných rezidentmi nepatriacimi do eurozóny, pričom sa využíva skutočnosť, že údajové požiadavky stanovené ECB sú plne v súlade s účasťou na koordinovanom prieskume portfóliových investícií, ktorý uskutočňuje MMF. Portfóliové investície v sebe nesú informáciu o investičných cenných papieroch, resp. majetkových cenných papieroch a dlhových cenných papieroch, ktoré vystupujú v podobe dlhopisov a úpisov a nástrojov peňažného trhu v prípade, že nie sú zaradené medzi priame investície alebo rezervné aktíva.

### ***Finančné deriváty***

Finančné derivát predstavujú finančné nástroje viazané na konkrétny finančný nástroj, ukazovateľ alebo komoditu, prostredníctvom ktorých sa dá na finančných trhoch obchodovať s konkrétnymi finančnými rizikami. Obchody s finančnými derivátmi sa považujú za samostatné transakcie a nie za súčasť hodnoty podkladových obchodov, na ktoré sa môžu viazať. Ide o transakcie a pozície v opciách, futures, swapoch, termínovaných devízových obchodoch, úverových derivátoch a zabudovaných derivátoch. Pod finančné deriváty by mali byť zaradené aj platby dodatočných záloh typu futures.

### ***Ostatné investície***

Ostatné investície sa definujú ako zvyšná kategória, ktorá obsahuje všetky finančné nástroje, ktoré nepatria medzi priame investície, portfóliové investície, finančné deriváty ani rezervné

---

<sup>1</sup> Charakteristika jednotlivých zložiek MIP vychádza z EURÓPSKA CENTRÁLNA BANKA USMERNENIE EURÓPSKEJ CENTRÁLNEJ BANKY z 2. mája 2003 o požiadavkách Európskej centrálnej banky na štatistické vykazovanie v oblasti štatistiky platobnej bilancie a medzinárodnej investičnej pozície a o výkaze medzinárodných rezerv (ECB/2003/7) (2003/381/ES)

aktíva. Medzi ostatné investície patria obchodné úvery, vklady peňazí a ostatné aktíva/ostatné pasíva. Zahŕňajú aj podvojnú položku za časovo rozlíšené príjmy z nástrojov zaradených medzi ostatné investície.

### 3 METODOLÓGIA

Analýzu vplyvu jednotlivých zložiek ukazovateľa MIP na hrubý domáci produkt realizujeme prostredníctvom ekonometrického modelu založenom na princípe modelu s korekčným členom<sup>2</sup>. Tento model je používaný pri určovaní miery vybočenia systému od tzv. ekvilibria a je založený na princípe kointegrácie časových radov<sup>3</sup>, pričom v jednej rovnici zachytáva tak dlhodobé, ako aj krátkodobé vzťahy medzi premennými.

Vzťah medzi premennými  $Y$  a  $X$  je možné vyjadriť prostredníctvom EC modelu v nasledovnom tvare:

$$\Delta Y_t = \alpha_0 + b_1 \Delta X_t - \pi u_{t-1} + u_t \quad (1)$$

kde

$$u_{t-1} = Y_{t-1} - \beta_1 - \beta_2 X_{t-1} \quad (2)$$

Model ECM je reprezentovaný vzťahom (1) pričom parameter  $b_1$  reprezentuje okamžitý dopad zmeny  $Y$  a  $X$ , parameter  $\pi$  predstavuje korekčný člen, ktorý hovorí o rýchlosti prispôsobenia sa krátkodobých vplyvov k dlhohodobej rovnováhe. Vzťah reprezentovaný rovnicou (2) je dlhodobým rovnovážnym stavom medzi premennými, ktoré sú nestacionárne.

Za hlavnú výhodu použitej metodiky ECM je možnosť zobrazenia tak krátkodobých, ako aj dlhodobých vzťahov. Premenné, s ktorými sa pracuje sú diferencované, čo spravidla rieši problém s nestacionaritou časových radov, a teda nedochádza k tzv. falošnej regresii.

### 4 DÁTA MODELU

Analýzu vplyvu skúmaných ukazovateľov priamych zahraničných investícií, portfóliových investícií, finančných derivátov a ostatných investícií na hrubý domáci produkt sme realizovali na štvrtročných dátach, pričom sme mali k dispozícii 51 pozorovaní, a to od 1Q2004-3Q2016. Údaje pre hrubý domáci produkt boli čerpané z databázy SLOVSTAT, údaje pre PZI, PORTI, FD a OSTI sme čerpali z databázy Národnej banky Slovenska. Všetky údaje sú v miliónoch EUR. Časové rady sú sezónne očistené v analýze označované skratkou „SA“. Časové rady boli zlogaritmované z dôvodu predpokladu zmenšenia rozptylu týchto časových radov, a teda riešenia problému ich nestacionarity. Zároveň premenné v logaritmickej tvare je možné interpretovať ako elasticity. Časové rady v logaritmickej tvare sú označované v analýze veľkým písmenom „L“. Všetky výpočty analýzy boli realizované prostredníctvom ekonometrického programu EViews.

### 5 VÝSLEDKY

Funkčný tvar modelu, ktorý sme analyzovali má nasledujúci tvar:

---

<sup>2</sup> Pre podrobnejšiu špecifikáciu modelu pozri [3]

<sup>3</sup> Pre podrobnejšiu špecifikáciu pozri [1]

$$HDP = f(PZI, PORTI, FD, OSTI) \quad (3)$$

pričom:

HDP – hrubý domáci produkt,  
PZI – priame zahraničné investície,  
PORTI – portfóliové investície,  
FD – finančné deriváty,  
OSTI – ostatné investície.

Vzhľadom na štatistickú nevýznamnosť ukazovateľa FD sme pri ďalšej práci s modelom od neho abstrahovali<sup>4</sup>. Pri tvorbe modelu sme postupovali systematicky v niekoľkých na seba nadväzujúcich krokoch. Prvým krokom bola analýza stacionarity pre všetky použité premenné. Stacionaritu sme testovali na 1 % hladine významnosti prostredníctvom Augmented Dickey-Fullerovho testu<sup>5</sup>. Z výsledkov, ktorých zhrnutie uvádzame v tabuľke 1 môžeme konštatovať, že všetky premenné sú integrované radu 1, a teda sú stacionárne na prvej diferencii.

**Tabuľka 1** Výsledky ADF testu pre premenné LHDP\_SA, LPZI\_SA, LPORTI\_SA a LOSTI\_SA

Premenná LHDP_SA	Level			1th difference		
	ADF test statistic	Test critical values	P - value	ADF test statistic	Test critical values	P - value
Intercept	-3,0552	-3,5744	0,0369	-4,3957	-3,5744	0,0010
Trend and Intercept	-2,6576	-4,1657	0,2582	-4,9697	-4,1611	0,0010
None	2,6326	-2,6140	0,9975	-3,2932	-2,6140	0,0015
Premenná LPZI_SA	Level			1th difference		
	ADF test statistic	Test critical values	P - value	ADF test statistic	Test critical values	P - value
Intercept	-3,1495	-3,5713	0,0294	-4,6300	-3,5744	0,0005
Trend and Intercept	-2,1163	-4,1567	0,5240	-6,5666	-4,1611	0,0000
None	5,2232	-2,6130	1,0000	-3,5859	-2,6140	0,0006
Premenná LPORTI_SA	Level			1th difference		
	ADF test statistic	Test critical values	P - value	ADF test statistic	Test critical values	P - value
Intercept	-0,7982	-3,5713	0,8107	-9,2596	-3,5744	0,0000
Trend and Intercept	-2,2828	-4,1567	0,4350	-9,1510	-4,1611	0,0000
None	2,3656	-2,6130	0,9951	-8,2308	-2,6140	0,0000
Premenná LOSTI_SA	Level			1th difference		
	ADF test statistic	Test critical values	P - value	ADF test statistic	Test critical values	P - value
Intercept	-2,5464	-3,5713	0,1111	-6,0857	-3,5744	0,0000
Trend and Intercept	-1,8586	-4,1567	0,6605	-6,3899	-4,1611	0,0000
None	1,5011	-2,6130	0,9654	-5,8800	-2,6140	0,0000

**Zdroj:** Vlastné spracovanie

Druhým krokom, ku ktorému sme následne pristúpili, bolo testovanie dlhodobých vzťahov, čím sme hľadali kointegračný vektor prostredníctvom Johansenovej kointegračnej analýzy<sup>6</sup>. Nami využívaný program EViews poskytuje dva kointegračné testy a to Trace test a Max-eigenvalue test. Výsledky z týchto testov, ktoré boli realizované na 5 % hladine významnosti uvádzame v tabuľke 2.

<sup>4</sup> Model (3) sme odhadli metódou najmenších štvorcov pre potvrdenie predpokladu o existencii vplyvu vybraných premenných na HDP, pričom ukazovateľ FD sa javil ako štatisticky nevýznamný.

<sup>5</sup> Pre bližšiu špecifikáciu pozri [1]

<sup>6</sup> Pre bližšiu špecifikáciu pozri [1]

**Tabuľka 2** Výsledky Johansenovej kointegračnej analýzy

HDP = f(PZI OSTI PORTI)	Nulové hypotézy	Trace Statistic	Critical Value 0,05	Max-Eigen Statistic	Critical Value 0,05
	r=0	57,028	47,856	36,037	27,584
	r<=1	20,99	29,797	11,155	21,131
	r<=2	9,835	15,494	9,069	14,264
	r<=3	0,765	3,841	0,765	3,841

**Zdroj:** Vlastné spracovanie

Výsledky poukazujú na to, že na hladine významnosti 5 % pri použití dvoch testov je v skúmanom modeli identifikovaný práve jeden kointegrujúci vektor, a teda môžeme zapísať rovnicu pre odhad dlhodobého vzťahu modelu prostredníctvom kointegračného vektora [1 -0,939 -0,032 0,079 -7,168] v tvare:

$$LHDP_{sa} = 7,168 + 0,939LPZI_{sa} + 0,032LPORIT_{sa} - 0,079LOSTI_{sa} \quad (4)$$

Zvolený kointegračný vektor v poslednom kroku použijeme na špecifikáciu samotného modelu s korekčným členom v nasledujúcom tvare:

$$\begin{aligned} D(LHDP\_SA) = & -0.0372*[LHDP\_SA(-1) - 0.9399*LPZI\_SA(-1) - \\ & 0.0325*LPORIT\_SA(-1) + 0.0791*LOSTI\_SA(-1) - 7.1684] + \\ & 0.5251*D(LHDP\_SA(-1)) + 0.1845*D(LHDP\_SA(-2)) - 0.1511*D(LPZI\_SA(-1)) + \\ & 0.3182*D(LPZI\_SA(-2)) - 0.0002*D(LPORIT\_SA(-1)) - 0.003*D(LPORIT\_SA(-2)) \\ & + 0.0285*D(LOSTI\_SA(-1)) + 0.005*D(LOSTI\_SA(-2)) - 0.0006 \end{aligned} \quad (5)$$

$$\begin{aligned} D(LPZI\_SA) = & 0.298*[LHDP\_SA(-1) - 0.9399*LPZI\_SA(-1) - \\ & 0.0325*LPORIT\_SA(-1) + 0.0791*LOSTI\_SA(-1) - 7.1684] - 0.416*D(LHDP\_SA(- \\ & 1)) - 0.335*D(LHDP\_SA(-2)) + 0.188*D(LPZI\_SA(-1)) - 0.116*D(LPZI\_SA(-2)) + \\ & 0.051*D(LPORIT\_SA(-1)) - 0.050*D(LPORIT\_SA(-2)) - 0.0677*D(LOSTI\_SA(-1)) \\ & - 0.059*D(LOSTI\_SA(-2)) + 0.028 \end{aligned} \quad (6)$$

$$\begin{aligned} D(LPORIT\_SA) = & -0.411*[LHDP\_SA(-1) - 0.9399*LPZI\_SA(-1) - \\ & 0.0325*LPORIT\_SA(-1) + 0.0791*LOSTI\_SA(-1) - 7.1684] + 0.192*D(LHDP\_SA(- \\ & 1)) - 0.451*D(LHDP\_SA(-2)) + 1.092*D(LPZI\_SA(-1)) + 0.474*D(LPZI\_SA(-2)) - \\ & 0.296*D(LPORIT\_SA(-1)) + 0.044*D(LPORIT\_SA(-2)) - 0.098*D(LOSTI\_SA(-1)) + \\ & 0.143*D(LOSTI\_SA(-2)) + 0.013 \end{aligned} \quad (7)$$

$$\begin{aligned} D(LOSTI\_SA) = & 0.675*[LHDP\_SA(-1) - 0.9399*LPZI\_SA(-1) - \\ & 0.0325*LPORIT\_SA(-1) + 0.0791*LOSTI\_SA(-1) - 7.1684] - 1.929*D(LHDP\_SA(- \\ & 1)) - 1.064*D(LHDP\_SA(-2)) - 0.164*D(LPZI\_SA(-1)) - 0.217*D(LPZI\_SA(-2)) - \\ & 0.008*D(LPORIT\_SA(-1)) - 0.114*D(LPORIT\_SA(-2)) - 0.184*D(LOSTI\_SA(-1)) + \\ & 0.034*D(LOSTI\_SA(-2)) + 0.067 \end{aligned} \quad (8)$$

Tvar modelu sme zvolili tak aby model v čo najväčšej miere spĺňal základné ekonometrické predpoklady, a teda aby v modeli nevystupovala heteroskedasticita ani autokorelácia, ktoré sme testovali prostredníctvom testov reziduálnych zložiek.

Z ekonomického hľadiska najvýraznejší pozitívny vplyv na ekonomiku sa preukázal pri premennej PZI, kde nárast tejto premennej o 1 % pri nemennosti ostatných premenných vyvoláva nárast HDP o 0,939 %. Pozitívny vplyv bol zaznamenaný aj pri ukazovateli PORTI, ktorý pri svojom zvýšení o 1 % a zachovaní ostatných premenných vyvoláva nárast HDP o 0,032 %. Negatívny vplyv na vývoj hrubého domáceho produktu Slovenska z dlhodobého hľadiska sme na základe výsledkov pozorovali pri ukazovateli OSTI. Pri zvýšení



ukazovateľa LOSTI\_sa o 1 % a nemennosti ostatných ukazovateľov dochádza k poklesu LHDP\_sa o 0,079 %.

Z výsledkov modelu VECM nevieme identifikovať jednoznačne existenciu krátkodobých vzťahov medzi vybranými premennými. Štatistická významnosť koeficientov prispôsobenia sa potvrdila len pre rovnice (5), (6), (8). Schopnosť korigovania krátkodobých odchýlok od rovnovážneho stavu len v rovnici (5). Koeficient prispôsobenia v tejto rovnici dosahuje hodnotu  $-0,0372$  %, čo nám hovorí o podiele korigovaných krátkodobých odchýlok.

## 6 ZÁVER

Aplikácia EC modelu, v ktorom sme sa zamerali na jednotlivé ukazovatele, z ktorých je tvorený ukazovateľ MIP si kládla za cieľ identifikovať tú zložku ukazovateľa, ktorá v najvýraznejšej miere vplýva na hrubý domáci produkt Slovenska. Týmto ukazovateľom s najvýraznejším vplyvom sú priame zahraničné investície, ktoré považujeme za veľmi dôležitý faktor rastu výkonnosti slovenskej ekonomiky a aj ako významný faktor jej ďalšieho rozvoja. Pozitívny vplyv ukazovateľa PZI môžeme vnímať cez jeho posilňovanie konkurenčného prostredia a následne stimulovanie podnikateľského prostredia.

Pri portfóliových investíciách sme tak výrazný pozitívny vplyv nezaznamenali čo môže byť spôsobené ich veľkou citlivosťou na vývoj v ekonomike, a teda investori môžu tieto investície vo veľmi krátkom čase stiahnuť z krajiny, čo môže v niektorých prípadoch vplývať na ekonomiku negatívne.

Negatívny vplyv ostatných investícií, ktorý sa prejavil na vývoji slovenského HDP môže byť rovnako pripisovaný ich veľkej citlivosti na vývoj v ekonomike vzhľadom na skutočnosť, že sú tvorené prevažne obchodnými úvermi, vkladmi peňazí, s ktorými môže ich investor svojvoľne disponovať<sup>7</sup>.

### Použitá literatúra

[1] ASTERIOU, D. – HALL, S. G. 2011. Applied econometrics. New York : Palgrave Macmillan, 2011. ISBN 987-0-230.27182-1

[2] DRITSAKI, M., DRITSAKI, C., ADAMOPOULOS, A. A. 2004. Causal Relationship between Trade, Foreign Direct Investment and Economic Growth for Greece. American Journal of Applied Sciences, vol. 1, No. 3, s. 230–235. ISSN 1546-9239.

[3] ENGLE, R. F. – GRANGER, C. W. J. 1987. *Cointegration and Error Correction: Representation, Estimation and Testing*. In *Econometrica*. 1987, vol. 55, no. 2, p. 251 – 276.

[4] EURÓPSKA CENTRÁLNA BANKA USMERNENIE EURÓPSKEJ CENTRÁLNEJ BANKY z 2. mája 2003 o požiadavkách Európskej centrálnej banky na štatistické vykazovanie v oblasti štatistiky platobnej bilancie a medzinárodnej investičnej pozície a o výkaze medzinárodných rezerv (ECB/2003/7) (2003/381/ES)

---

<sup>7</sup> Príspevok bol vypracovaný v rámci riešenia projektu APVV-14-0787 “Zladenie verejných financií a starobného dôchodkového zabezpečenia (Návrh udržateľného a rast podporujúceho dôchodkového systému pre starnúcu slovenskú ekonomiku)”

[5] JÁNOŠÍK, P. 2010. Medzinárodná investičná pozícia. Biatic: Národná banka Slovenska, ročník 18, 5/2010. dostupné na internete :

[www.nbs.sk/\\_img/Documents/PUBLIK/.../07\\_biatic0510\\_janosik.pdf](http://www.nbs.sk/_img/Documents/PUBLIK/.../07_biatic0510_janosik.pdf)

[6] KÖNIG, B. 2013. Analýza nezamestnanosti a jej makroekonomických dopadov. In: Medzinárodný vedecký seminár „Nové trendy v ekonometrii a operačnom výskume“. EKONÓM. Bratislava 2013. ISBN: 978-80-225-3786-5

[7] LUKÁČIKOVÁ, A. – LUKÁČIK, M. – SZOMOLÁNYI, K. 2007. Ekonometrické prognózovanie importu Slovenskej Republiky na základe modelov s korekčným členom. In: Ekonomické rozhľady. Ekonomická univerzita v Bratislave. Bratislava. 2/2007. Ročník XXXVI

[8] NÁRODNÁ BANKA SLOVENSKA. 2006. Priame zahraničné investície. Bratislava. 2006. Dostupné na internete:

[http://www.nbs.sk/\\_img/Documents/STATIST/SPB/PZI/PZI\\_2006.pdf](http://www.nbs.sk/_img/Documents/STATIST/SPB/PZI/PZI_2006.pdf)

#### **Kontaktné údaje**

Ing. Miroslava Jánošová, PhD.

Ekonomický ústav SAV

Šancová 56

811 05 Bratislava

email: [mdolinajcova@gmail.com](mailto:mdolinajcova@gmail.com)

# POTENCIÁLNE KOALÍCIE PO PARLAMENTNÝCH VOĽBÁCH V ROKOCH 2012 A 2016 V SLOVENSKEJ REPUBLIKE<sup>1</sup>

## POSSIBLE COALITIONS IN THE SLOVAK REPUBLIC AFTER GENERAL ELECTIONS 2012 AND 2016

*Rudolf Kucharčík, Zuzana Čičková*

### **Abstrakt**

Význam koalícií politických strán závisí od výsledkov parlamentných volieb. V prípade, že niektorá politická strana získa väčšinu poslaneckých mandátov v parlamente, stáva sa rozhodujúcim aktérom na politickej scéne (prípád v Slovenskej republike v roku 2012) a akékoľvek koalície majú len zanedbateľný politický vplyv. Na strane druhej, ak je parlamentná väčšina výsledkom koalíčnej dohody (prípád v Slovenskej republike v roku 2016), narastá politický vplyv aj politickej strany s menším (malým) počtom mandátov v prípade, že sa stáva súčasťou vládnej koalície. Teória hier poskytuje mnohé nástroje umožňujúce kvantifikovať vplyv politických strán. V príspevku sa tiež zaoberáme vyčíslením sily politických strán na základe Shapleyho–Shubikovo ukazovateľa a Banzhafovo ukazovateľa.

*Keľúčové slová: teória hier, voľby, koalície, politické strany*

### **Abstract**

The importance of political parties' coalitions depends on the results of parliamentary elections. If some political party gains majority of the seats in the parliament, it becomes decisive player on the political scene (the case of the Slovak republic in 2012). Other coalitions have only little political influence. On the other hand, if the majority in the parliament is the result of coalition agreement (the case of the Slovak republic in 2016), the influence of the small political party is growing if it is part of governmental coalition. Game theory provides many tools to quantify the impact of political parties. In the paper we also deal with the quantification of the power of political parties based on the Shapley-Shubik power index and Banzhaf power index.

*Keywords: game theory, election, coalitions, political parties*

## **1 ÚVOD**

Z hľadiska existencie politických koalície nastali v Slovenskej republike po parlamentných voľbách v roku 2012 a v roku 2016 dve diametrálne odlišné situácie (pozri grafy 1-5).

V roku 2012 sa rozhodujúcim politickým subjektom stala politická strana Smer-SD (sociálna demokracie). Získala presvedčivú väčšinu mandátov (83) v 150 členom parlamente. Žiadna z ďalších politických strán nebola schopná iniciovať samostatne ani zvolanie mimoriadnej schôdze parlamentu, na čo je potrebná podpora minimálne 20 poslancov (bližšie pozri Ústavu SR). Išlo o politické strany Kresťansko-demokratické hnutie (KDH), Sloboda a Solidarita (SaS), Obyčajní ľudia a nezávislé osobnosti (OĽaNO), Most-Híd, Slovenská demokratická a kresťanská únia – Demokratická stranu (SDKÚ-DS).

---

<sup>1</sup> Príspevok vznikol ako súčasť výskumnej úlohy VEGA 1/0351/17 Aplikácia vybraných modelov teórie hier pri riešení niektorých ekonomických problémov Slovenska

Po voľbách v roku 2016 sa žiadna z politických strán na parlamentnú väčšinu spoliehať nemohla. Dokonca nebolo možné vytvoriť ani funkčnú koalíciu zloženú z dvoch politických strán. Parlamentná väčšina mohla byť vytvorená minimálne z troch strán – za predpokladu, že súčasťou tejto koalície bude aj víťaz parlamentných volieb strana Smer-SD.

Príspevok sa bude zaoberať teoretickým možnostiam spojeným s vznikom koalícií bezprostredne po parlamentných voľbách (o teórii koalícií, volieb, resp. teórii hier pozri aj: Cabada-Charvát-Stulík, 2015, Démuth 2013, Krejčí 2007, Drulák 2010, Novák-Lebeda a kol., 2004, Novák a kol., 2011). V príspevku nebudeme analyzovať politickú situáciu (napr. rozpady politických strán alebo odchody poslancov počas volebného obdobia, resp. vylúčenie spolupráce s niektorou z politických strán ostatnými politickými subjektmi).

Teória hier ponúka nástroje umožňujúce vyčíslieť silu hráčov (politických strán) v koalícii. Budeme uvažovať s týmito predpokladmi:

- poslanci jednej strany hlasujú jednotne,
- ak sa vytvorí koalícia strán, tiež všetci jej členovia hlasujú jednotne,
- možno vytvoriť ľubovoľnú koalíciu.

V príspevku vyčíslíme dva ukazovatele Sharpleyho - Shubikov ukazovateľ, ktorý udáva pravdepodobnosť toho, že strana je nevyhnutná pri zostavovaní väčšinových koalícií (všetkých možných) a Banzhafov ukazovateľ, ktorý ukazuje blokovaciu silu hráča v procese hlasovania (Dlouhý-Fiala,2015).

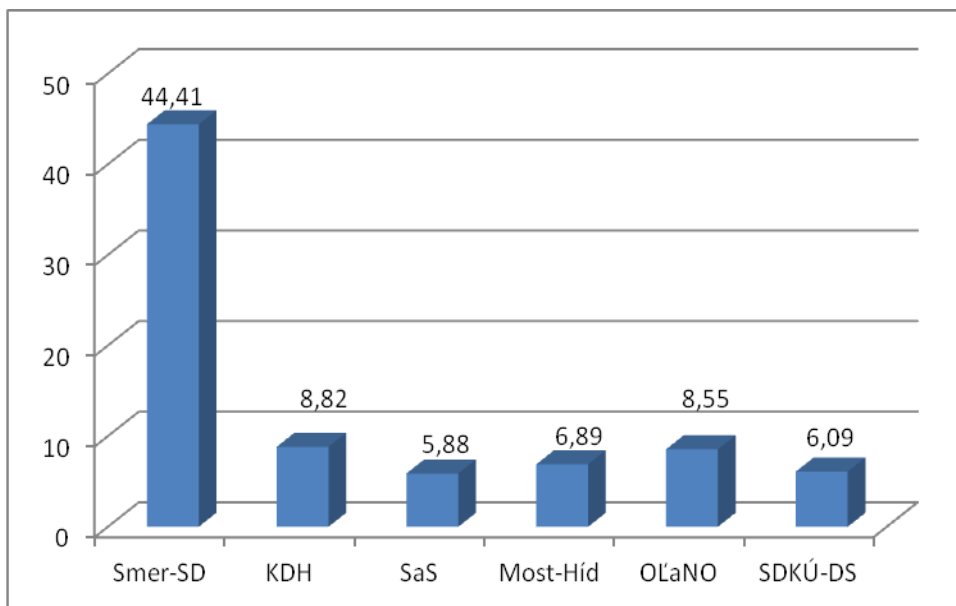
## **2 KOALÍCIE PO PARLAMENTNÝCH VOĽBÁCH V ROKU 2012**

Celkový počet možných koalícií po parlamentných voľbách v roku 2012 dosiahol číslo 63. Z politického hľadiska však zásadnejší význam nemali. Dominantným politickým subjektom sa stala strana vedená Robertom Ficom Smer-SD. Počet mandátov 83 bol dostatočný na vytvorenie tzv. jednofarebnej vlády. V minulosti osvedčená stranícka disciplína strany Smer-SD bola dostatočným predpokladom politickej stability na ďalšie štyri roky.

Koalície mali väčší význam len pri prijímaní ústavných zákonov (potrebný počet poslancov na prijatie ústavného zákona, resp. na zmenu ústavy bol 90). Sila politickej strany Smer-SD bola dostatočná na to, aby jej na zmenu tohto typu stačil len jeden partner. Strana Smer-SD mohla takto vyjednať päť rozdielnych koalícií. Ďalšími možnosťami bol vznik trojkoalícií (celkovo 10), štvorkoalícií (celkovo 10), päťkoalícií (celkovo 5) alebo jednej šesťkoalícií. Samozrejme tieto prípady (aj keď formálne neboli potrebné ani na prijatie ústavného zákona, resp. zmenu ústavy) by prípadnému rozhodnutiu (zmena ústavy, prijatie ústavného zákona) dali väčšiu legitimitu.

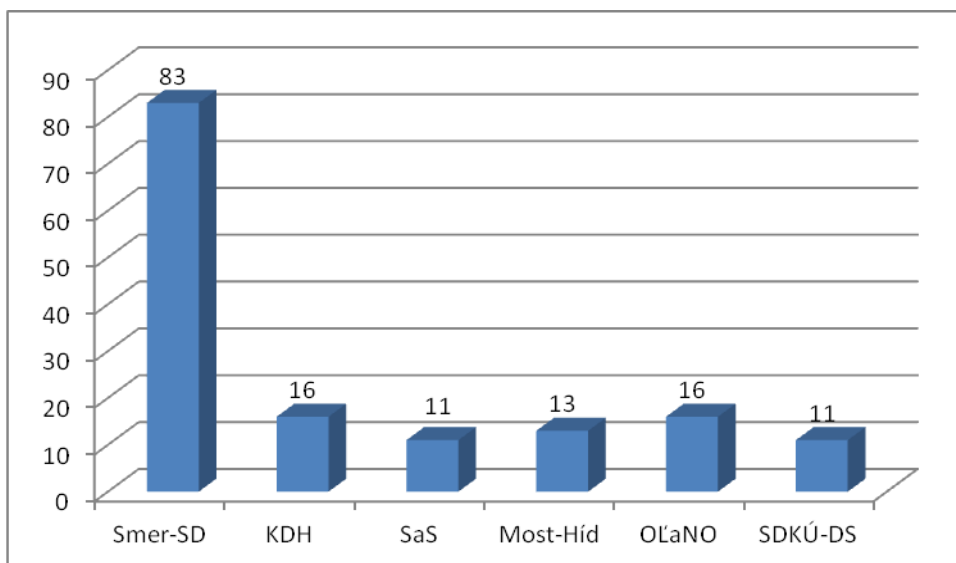
Druhou situáciou, v ktorej mohli koalície zohrať výraznejšiu úlohu, bola parlamentná činnosť. V zmysle Ústavy Slovenskej republiky je napr. potrebný počet 15 poslancov na súhlas s kandidatúrou na post prezidenta SR. Svojho kandidáta tak mohli navrhnúť Smer-SD, KDH prípadne OĽaNO. Zvyšné tri politické strany potrebovali na takýto krok najmenej jedného ďalšieho politického partnera. Ešte komplikovanejšia bola pre opozičné politické strany situácia, keď sa usilovali o zvolanie mimoriadnej schôdze NR SR (napr. s cieľom rokovat' o vyslovení nedôvery členovi vlády SR). Žiadna z opozičných politických strán by nebola schopná vyvolať takéto rokovanie samostatne, keďže je nevyhnutná iniciatíva minimálne 20 poslancov. Každá opozičná strana tak potrebovala minimálne jedného partnera.

**Graf 1 Výsledky volieb do NR SR v roku 2012 (v %)**



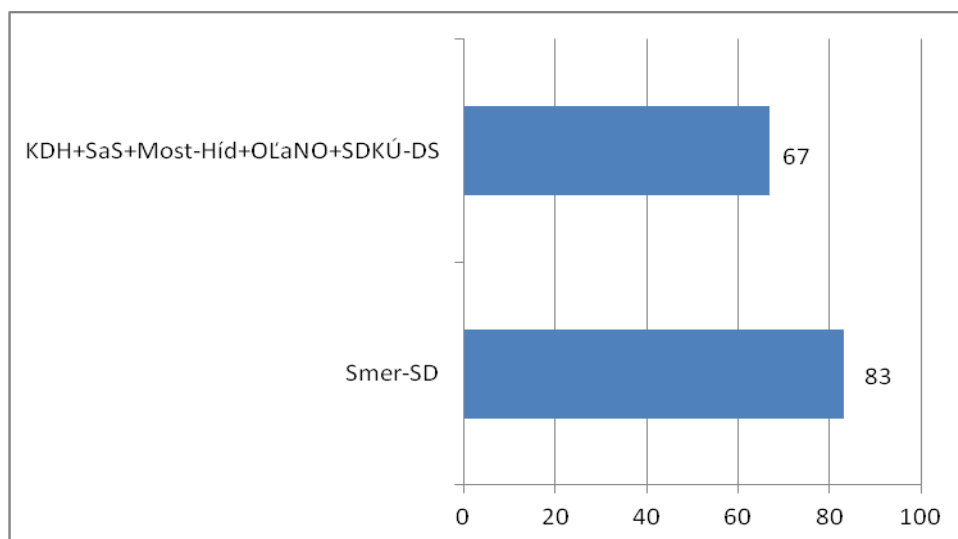
Zdroj: Štatistický úrad SR

**Graf 2 Výsledky volieb do NR SR v roku 2012 (počet mandátov)**



Zdroj: Štatistický úrad SR

**Graf 3 Sila strany Smer-SD vs ostatné politické strany po voľbách v roku 2012**



Zdroj: Štatistický úrad SR

Berúc do úvahy Banzhafov a Sharpleyho-Shubikov ukazovateľ, dostávame v roku 2012 nasledujúce hodnoty:

**Tabuľka 1**

Banzhafov ukazovateľ - konštanta 75		Banzhafov ukazovateľ - konštanta 90	
SMER-SD	1	SMER-SD	0.8611
KDH	0	KDH	0.0278
SAS	0	SAS	0.0278
MOST-HÍD	0	MOST-HÍD	0.0278
OLaNO	0	OLaNO	0.0278
SDKÚ-DS	0	SDKÚ-DS	0.0278

**Tabuľka 2**

Sharpleyho - Shubikov ukazovateľ - 75		Sharpleyho - Shubikov ukazovateľ - 90	
SMER-SD	1	SMER-SD	0.8333
KDH	0	KDH	0.0333
SaS	0	SaS	0.0333
MOST	0	MOST	0.0333
OLaNO	0	OLaNO	0.0333
SDKÚ-DS	0	SDKÚ-DS	0.0333

### 3 KOALÍCIE PO PARLAMENTNÝCH VOĽBÁCH V ROKU 2012

Situácia po parlamentných voľbách v roku 2016 bola z hľadiska koalícií zložitejšia. Celkový počet koalícií dosiahol číslo 255 (dvoj až osemkoalícií). Žiadna z politických strán nebola schopná vytvoriť jednofarebnú vládu. Na vytvorenie funkčnej koalície však nestačilo ani spojenie dvoch politických strán.

Nová vláda (ak sa mala opierať o parlamentnú väčšinu) potrebovala podporu minimálne troch politických strán (súčasťou každej však musela byť politická strana Smer-SD). Teoreticky takýchto koalícií mohlo vzniknúť 12 (reálne maximálne 9, nakoľko môžeme predpokladať, že žiadna z politických strán by nebola ochotná vytvoriť spoločnú vládu s politickou stranou Ľudová strana – Naše Slovensko). Žiadna z možných trojkoalícií by nedisponovala ústavnou väčšinou.

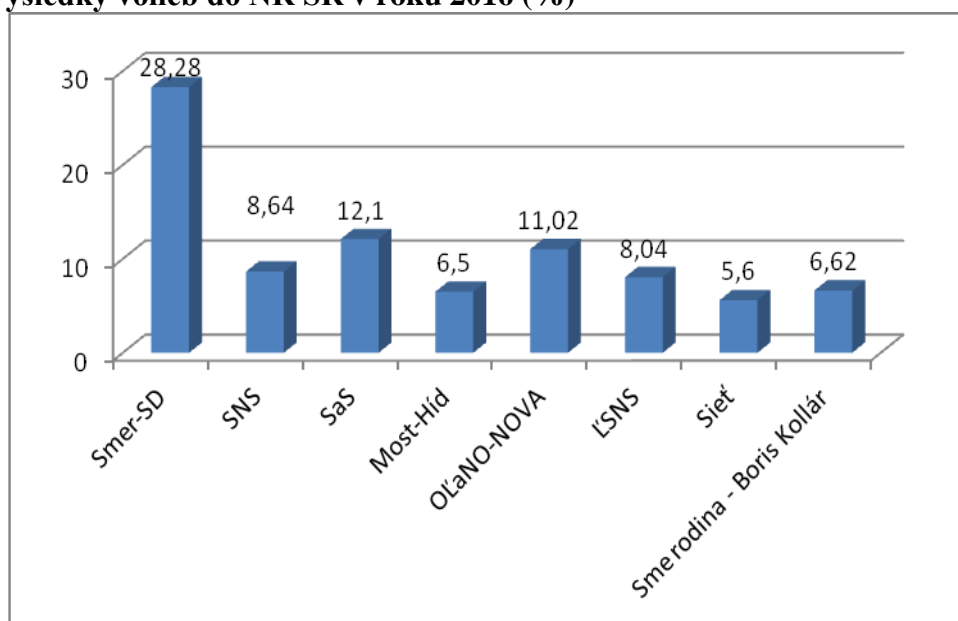
Počet možných (teoretických) funkčných (t.j. s parlamentnou väčšinou) štvorkoalícií bol 35. Aj v tomto prípade však vládnu stranou musela byť strana Smer-SD. Štvorkoalícia, ktorá by vynechala stranu Smer-SD, by parlamentnou väčšinou nedisponovala. Celkovo 23 rôznych štvorkoalícií by disponovalo 90 poslancami, t.j. ústavnou väčšinou.

Čo sa týka funkčných päťkoalícií, tých mohlo po parlamentných voľbách v roku 2016 vzniknúť 42. Z tohto počtu by 37 disponovalo ústavnou väčšinou (súčasťou každej by však musela byť politická strana Smer-SD). Celkovo 5 koalícií by disponovalo len parlamentnou väčšinou (súčasťou žiadnej z nich by nebola politická strana Smer-SD).

Počet rôznych funkčných šesťkoalícií bol po parlamentných voľbách v roku 2016 28. Z nich len štyri nedisponovali ústavnou väčšinou. Dve z nich disponovali ústavnou väčšinou aj bez účasti politickej strany Smer-SD. Sedemkoalícií mohlo vzniknúť celkovo 8. Osemkoalícia jedna. Tieto prípady už zaručovali ústavnú väčšinu vždy.

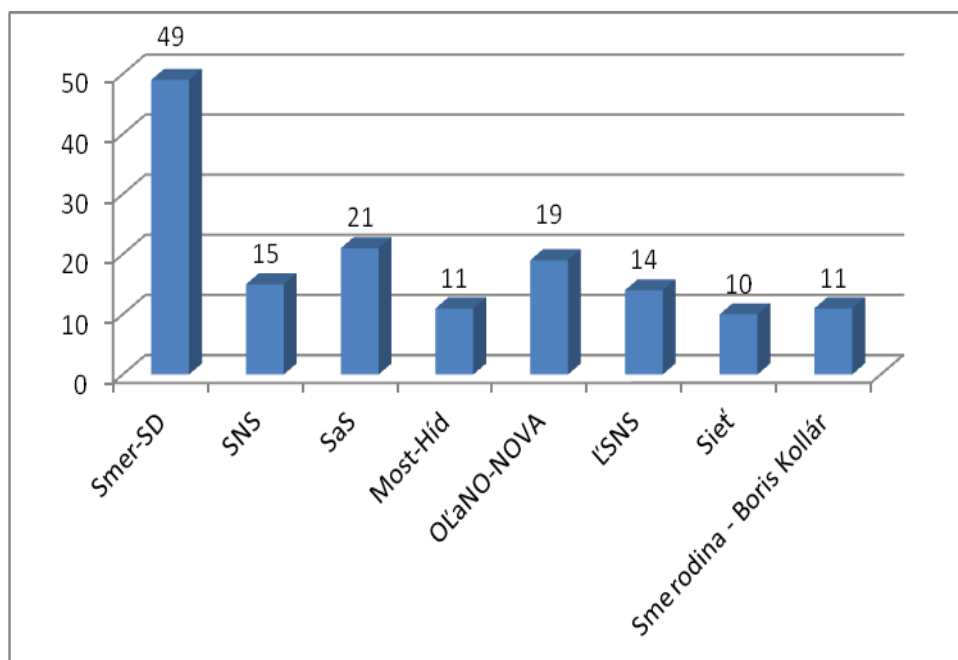
Vznik koalícií predpokladala aj bežná parlamentná činnosť – len Smer-SD a SaS disponovali dostatočným počtom poslancov na zvolanie mimoriadnej schôdze parlamentu. Ostatné politické strany potrebovali na takúto iniciatívu partnera (vždy stačil jeden). Most-Híd, koalícia OĽaNO-NOVA, ĽSNS, Siet' a Sme rodina – Boris Kollár by potrebovali minimálne jedného koalíčného partnera aj na prípadnú nomináciu kandidáta do prezidentských volieb.

**Graf 4 Výsledky volieb do NR SR v roku 2016 (%)**



Zdroj: Štatistický úrad SR

**Graf 5 Výsledky volieb do NR SR v roku 2016 (počet mandátov)**



Zdroj: Štatistický úrad SR

Berúc do úvahy Banzhafov a Sharpleyho-Shubikov ukazovateľ, dostávame v roku 2016 nasledujúce hodnoty:

**Tabuľka 3 Banzhafov ukazovateľ za rok 2016**

Banzhafov ukazovateľ - konštanta 75		Banzhafov ukazovateľ - konštanta 90	
SMER-SD	0.4211	SMER-SD	0.3925
SNS	0.0877	SNS	0.0841
SaS	0.1228	SaS	0.1402
MOST-HÍD	0.0614	MOST-HÍD	0.0654
OĽaNO - NOVA	0.1228	OĽaNO - NOVA	0.1121
ĽSNS	0.0702	ĽSNS	0.0841
SIEŤ	0.0526	SIEŤ	0.0561
SME-BK	0.0614	SME-BK	0.0654



**Tabuľka 4 Sharpleyho – Shubikov ukazovateľ za rok 2016**

Sharpleyho - Shubikov ukazovateľ - konštant 75		Sharpleyho - Shubikov ukazovateľ - konštant 90	
SMER-SD	0.4048	SMER-SD	0.4357
SNS	0.0905	SNS	0.0810
SaS	0.1286	SaS	0.1238
MOST-HÍD	0.0619	MOST-HÍD	0.0667
OĽaNO-NOVA	0.1286	OĽaNO-NOVA	0.1024
ĽSNS	0.0714	ĽSNS	0.0810
SIEŤ	0.0524	SIEŤ	0.0429
SME-BK	0.0619	SME-BK	0.0667

#### 4 ZÁVER

Slovenská republika je parlamentnou demokraciou, čo znamená, že na rozdiel od prezidentských systémov je vláda závislá od podpory parlamentu. V roku 2012 nastala prvý krát od roku 1989 situácia, keď výsledkom parlamentných volieb nebola koalíčná vláda. Aj keď táto situácia je výhodná najmä z pohľadu stability politického systému, nie je možné predpokladať, že sa s ňou vzhľadom na štiepne línie slovenskej spoločnosti budeme stretávať pravidelnejšie. Predpokladať môžeme skôr potrebu vytvárania koalícií. Otáznym je počet politických strán, od podpory ktorých bude vláda závislá a následné alternatívy, ktoré sa počas politických diskusií politických lídrov budú ponúkať. Na analýzu politickej situácie možno využívať rôzne nástroje teórie hier a analyzovať hlasovacie a koalíčné hry. V príspevku sme prezentovali vyčíslenie Sharpleyho - Shubikovho ukazovateľa a Banzhafovho ukazovateľa pre politickú scénu Slovenska.

**Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/0351/17 Aplikácia vybraných modelov teórie hier pri riešení niektorých ekonomických problémov Slovenska**

#### Použitá literatúra

1. BALÁŽ, V. 2014. Komplexné voľby. Veda, Bratislava, 2014.
2. CABADA, L. – CHARVÁT, J. – STULÍK, O. 2015., Vydavatelství a nakladatelství Aleš Čeněk, s.r.o., Praha – Plzeň, 2015.
3. DÉMUTH, A. 2013. Teória hier a problém rozhodovania. Filozofická fakulta Trnavskej univerzity v Trnave, Trnava, 2013.
4. DLOUHÝ, M. – FIALA, P. 2015. Teorie ekonomických a politických her. Oeconomica, Praha, 2015.
5. DRULÁK, P. 2010. Teorie mezinárodních vztahů. Portál, Praha, 2010.
6. KREJČÍ, O. 2007. Mezinárodní politika. Ekopress, Praha, 2007.
7. NOVÁK, M. et al. 2011. Úvod do studia politiky. Slon, Praha, 2011.
8. NOVÁK, M. – LEBEDA, T. 2004. Volební a stranické systémy. ČR v mezinárodním srovnání. Aleš Čeněk, Dobrá Voda u Pelhřimova, 2004.
9. Štatistický úrad Slovenskej republiky. [online]. [Citované 1. 11. 2017.] Dostupné na internete: [www.statistics.sk](http://www.statistics.sk)

10. Ústava Slovenskej republiky. [online]. [Citované 1. 11. 2017.] Dostupné na internete: [https://www.nrsr.sk/web/Static/sk-SK/NRSR/Doc/zd\\_ustava\\_2012.pdf](https://www.nrsr.sk/web/Static/sk-SK/NRSR/Doc/zd_ustava_2012.pdf)

### **Kontaktné údaje**

PhDr. Rudolf Kucharčík, PhD.  
Ekonomická univerzita v Bratislave, Fakulta medzinárodných vzťahov  
Dolnozemska 1/b, 852 35 Bratislava  
Tel: (421 2) 67 295 470  
email: [rudolf.kucharcik@euba.sk](mailto:rudolf.kucharcik@euba.sk)

doc. Ing. Zuzana Čičková, PhD.  
Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky  
Dolnozemska 1/b, 852 35 Bratislava  
Tel: (421 2) 67 295 820  
email: [zuzana.cickova@euba.sk](mailto:zuzana.cickova@euba.sk)

# VYUŽITIE METÓDY MAXIMÁLNEJ VIEROHODNOSTI V PROGRAME EViews<sup>1</sup>

## USING OF MAXIMUM LIKELIHOOD METHOD IN EViews

*Martin Lukáčik*

### Abstrakt

Rozvoj ekonometrickej analýzy je v posledných rokoch úzko spojený s rozvojom výpočtovej techniky a zodpovedajúceho softvéru. Bez kvalitného programového vybavenia by sme nemohli realizovať takmer žiadne analýzy. V tomto príspevku sme sa rozhodli predstaviť, akým spôsobom sa dá v programe EViews využiť metóda maximálnej vierohodnosti. Metódu prezentujeme na príklade odhadu nelineárneho modelu, ktorým je produkčná funkcia s konštantnou elasticitou substitúcie, ktorá je už dlhšie predmetom nášho záujmu.

*Keľúčové slová: ekonometrická analýza, metóda maximálnej vierohodnosti, EViews*

### Abstract

The development of econometric analysis is in recent years closely linked to the development of computers and appropriate software. Without quality software, we could not implement almost any analysis. In this paper, we decided to introduce how the maximum likelihood method can be used in EViews. We present the method on an example of nonlinear model of constant elasticity of substitution production function, which is a feature that has long been our concern.

*Keywords: econometric analysis, maximum likelihood method, EViews*

## 1 ÚVOD

Okrem metódy najmenších štvorcov, ktorá je najpoužívanjšou metódou odhadu neznámych parametrov, existujú ďalšie metódy odhadu parametrov. Najznámejšou takou metódou je metóda maximálnej vierohodnosti (Hatrák, 2007). Je postavená na maximalizácii logaritmu vierohodnostnej funkcie, ktorá je súčinom funkcií hustoty pravdepodobnosti nezávislých náhodných premenných výberového súboru. Pre združené rozdelenie pravdepodobnosti výberu platí, že je súčinom individuálnych rozdelení premenných:

$$f(y_1, y_2, \dots, y_n) = f(y_1)f(y_2) \dots f(y_n) \quad (1)$$

ak sú premenné nezávislé. A ak majú normálne rozdelenie, tak pre  $f(y_i)$  platí:

$$f(y_i) = (2\pi\sigma^2)^{-1/2} e^{-\left(\frac{1}{2} \frac{(y_i - \mu)^2}{\sigma^2}\right)} \quad (2)$$

kde  $\sigma^2$  predstavuje rozptyl a  $\mu$  je priemer rozdelenia náhodnej premennej  $Y$ . Podľa konvencie zapisujeme  $Y \sim N(\mu, \sigma^2)$ . Vierohodnostná funkcia je totožná s funkciou združeného rozdelenia pravdepodobnosti výberu, pričom je chápaná ako funkcia hodnôt parametrov pri fixovaných

---

<sup>1</sup> Príspevok vznikol s podporou projektu VEGA 1/0444/15 "Ekonometrická analýza produkčných možností ekonomiky a trhu práce na Slovensku".

hodnotách výberu ( $y_1, y_2, \dots, y_n$  poznáme, ale nepoznáme parametre). Pomocou nej sa za daných podmienok dajú určiť charakteristiky vybraného rozdelenia, ktoré generujú výberové dáta s najväčšou pravdepodobnosťou.

Vierohodnostná funkcia  $L$  modelu pri identickom a nezávislom normálnom rozdelení všetkých prvkov vektora náhodnej zložky má tvar (Lukáčik et al, 2013):

$$L(\beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2\right)} \quad (3)$$

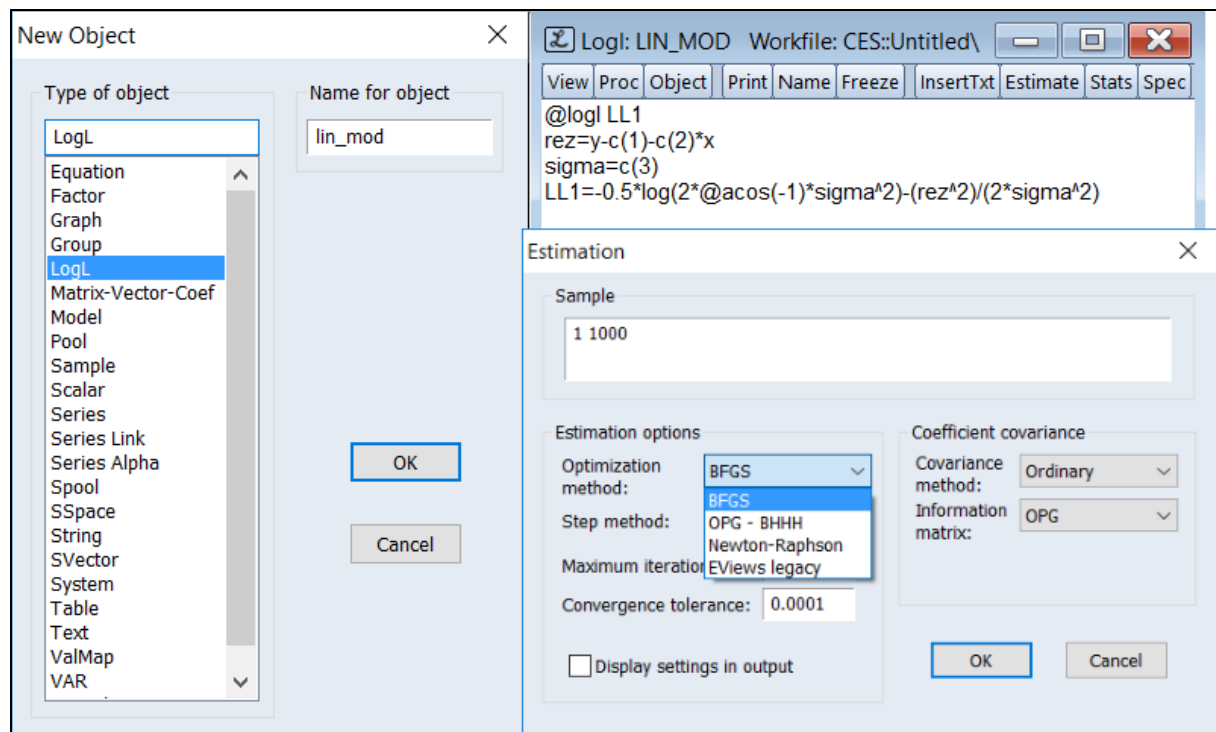
Aby sme nemuseli hľadať maximum funkcie (3), úloha sa zjednodušuje. Využijeme pritom logaritmickú transformáciu a jej vlastnosť, že ponecháva extrémny v rovnakých bodoch ako pôvodná funkcia. Riešime úlohu:

$$\max_{\beta, \sigma^2} \ln L = \max_{\beta, \sigma^2} \ln \left[ (2\pi\sigma^2)^{-n/2} e^{\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2\right)} \right] = \max_{\beta, \sigma^2} \left[ -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(\mathbf{x}_i))^2 \right] \quad (4)$$

v ktorej z prvých parciálnych derivácií položených za rovnajúce sa nule získame sústavu rovníc, ktorých riešením sú estimátory metódy maximálnej vierohodnosti parametrov modelu.

## 2 MAXIMÁLNA VIEROHODNOSŤ V PROGRAME EViews

Od verzie programu EViews 9 tvorcovia aplikujú metódu maximálnej vierohodnosti na odhad parametrov modelov, ktoré obsahujú autoregresné členy  $AR(p)$  alebo členy kľzavých priemerov  $MA(q)$ . Ale už v skorších verziách bolo umožnené vytvorenie objektu typu  $LogL$  teda  $log$  likelihood, ktorý umožňoval odhad parametrov metódou maximálnej vierohodnosti.



Obrázok 1: Odhad parametrov metódou maximálnej vierohodnosti v programe EViews

Ak by sme predpokladali najjednoduchší lineárny model v tvare:

$$y_i = f(x_i) = \beta_0 + \beta_1 x_i + u_i \quad (5)$$

Tak po výbere nového objektu typu *LogL* by sme zadali špecifikáciu (upravený Giles, 2013):

```
@logl LL1
rez=y-c(1)-c(2)*x
sigma=c(3)
LL1=-0.5*log(2*@acos(-1)*sigma^2)-(rez^2)/(2*sigma^2)
```

Po zvolení výberu (*sample*) získame odhady parametrov  $\beta_0$ ,  $\beta_1$  a  $\sigma$  v programom zobrazených hodnotách C(1), C(2) a C(3). Špecifikáciu sme vytvorili tak, že v prvom riadku je názov premennej, kde sa vypočítava príspevok *i*-tého pozorovania k logaritmu vierohodnosti, čo

vidíme v štvrtom riadku, ktorý zodpovedá vzťahu  $-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(rez_i)^2}{2\sigma^2}$ . V ňom sú použité

v druhom a treťom riadku zadané parametre, ktoré špecifikujú vzťah pre reziduál a definujú štandardnú odchýlku. Funkcia @acos(-1) reprezentuje čo najpresnejšie číslo  $\pi$ .

Mohli by sme využiť aj zabudovanú funkciu EViewsu @dnorm pre hustotu štandardného normálneho rozdelenia a zadať špecifikáciu podľa sprievodcu programu (pozri User's Guide):

```
@logl LL1
rez=y-c(1)-c(2)*x
sigma=c(3)
LL1=log(@dnorm(rez/sigma))-log(sigma^2)/2
```

V prípade lineárneho modelu s autoregresnou náhodnou zložkou 1. rádu v tvare:

$$y_t = f(x_t) = \beta_0 + \beta_1 x_t + u_t, \text{ kde } u_t = \rho u_{t-1} + \varepsilon_t \quad (6)$$

potrebujeme vytvoriť pomocnú umelú premennú napr. *d1*, ktorá sa bude rovnať 1 v prvom pozorovaní a inak 0. Dôvodom na jej vytvorenie je využitie funkcie @recode, ktorá reprezentuje klasickú funkciu *if*, teda priradenie druhého argumentu ak je prvý pravdivý, inak sa priradí tretí argument. Takýmto spôsobom riešime stratu prvého pozorovania.

Pri zadaní vierohodnosti využijeme reziduál modelu zovšeobecnených diferencií, ktorý sa získa odpočítaním  $\rho$  násobku o jedno obdobie oneskoreného modelu (6) od pôvodného modelu (6). Model zovšeobecnených diferencií má tvar:

$$y_t = \beta_0(1-\rho) + \rho y_{t-1} + \beta_1 x_t - \rho \beta_1 x_{t-1} + \varepsilon_t \quad (7)$$

Po výbere nového objektu typu *LogL* by sme zadali špecifikáciu:

```
@logl LL1
var = @recode(d1=1,s2(1)/(1-ro(1)^2),s2(1))
rez = @recode(d1=1,y-c(1)-c(2)*x,y-c(1)*(1-ro(1))-ro(1)*y(-1)-c(2)*x+ro(1)*c(2)*x(-1))
srez = rez/@sqrt(var)
LL1 = log(@dnorm(srez))-log(var)/2
```

Alebo od verzie EViews 9 stačí zadať odhad rovnice so zadaným autoregresným členom:

```
equation rov1.ls y c x ar(1)
```

pričom odhady parametrov  $\beta_0$ ,  $\beta_1$  sú zobrazené v hodnotách C(1) resp. C a C(2) resp. X, odhad  $\rho$  je zobrazený v hodnotách RO(1) resp. AR(1) a odhad rozptylu  $\sigma^2$  je zobrazený v hodnotách S2(1) resp. SIGMASQ.

Metóda maximálnej vierohodnosti sa používa aj pri odhade modelov s kvalitatívnou závislou premennou ako je logit (pozri Lukáčiková, 2011) alebo probit (pozri Greene, 2011).

Pravdepodobnostný model logit využívajúci logistickú distribučnú funkciu:

$$p_i = p(y_i^* = 1 | \mathbf{x}_i) = \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}}}{1 + e^{\mathbf{x}_i^T \boldsymbol{\beta}}} \quad (8)$$

pri odvodení maximálnej vierohodnosti maximalizuje funkciu:

$$\max_{\boldsymbol{\beta}} \ln L = \max_{\boldsymbol{\beta}} \left[ \sum_{i=1}^n \left( y_i \ln \left( \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}}}{1 + e^{\mathbf{x}_i^T \boldsymbol{\beta}}} \right) + (1 - y_i) \ln \left( \frac{1}{1 + e^{\mathbf{x}_i^T \boldsymbol{\beta}}} \right) \right) \right] \quad (9)$$

Ak by sme predpokladali, že závislá premenná je označená  $y$  a máme napr. dve vysvetľujúce premenné  $x_1$  a  $x_2$ , po výbere nového objektu typu *LogL* by sme zadali špecifikáciu:

```
@logl LL1
xb = -b(1)-b(2)*x1-b(3)*x2
pr = 1/(1+exp(xb))
LL1 = y*log(pr)+(1-y)*log(1-pr)
```

Rovnaký výsledok, ale s väčším množstvom sprievodných štatistík získame odhadom rovnice príkazom EViewsu:

```
equation rov1.binary(d=1) y c x1 x2
```

### 3 CES PRODUKČNÁ FUNKCIA

Produkčné funkcie všeobecne a aj produkčná funkcia s konštantnou elasticitou substitúcie, sú už dlhšie predmetom záujmu kolektívu autorov Szomolányi, Lukáčik a Lukáčiková. Špecifiká produkčných funkcií ako časové hľadisko aj rôzne spôsoby ich odhadu boli prezentované vo viacerých príspevkoch (Szomolányi, 2014; Szomolányi, 2017 a Szomolányi et al, 2017).

Produkčná funkcia s konštantnou elasticitou substitúcie (CES) má tvar (Lukáčik, 2013):

$$Q = \gamma \left[ \delta K^{-\rho} + (1 - \delta) L^{-\rho} \right]^{-\nu/\rho} \quad \gamma > 0, \nu > 0, 0 < \delta < 1, -1 < \rho < \infty, \rho \neq 0 \quad (10)$$

Parameter efektívnosti  $\gamma$  má rovnakú úlohu ako pri Cobbovej Douglasovej produkčnej funkcii. Kladný má byť preto, aby nárast faktorov viedol k nárastu produkcie, uvažuje sa aj rovný 1.

Stupeň homogénnosti  $\nu$  má na starosti výnosy z rozsahu, takže platí známe:

$$\gamma \left[ \delta (\lambda K)^{-\rho} + (1 - \delta) (\lambda L)^{-\rho} \right]^{-\nu/\rho} = \gamma (\lambda^{-\rho})^{-\nu/\rho} \left[ \delta K^{-\rho} + (1 - \delta) L^{-\rho} \right]^{-\nu/\rho} = \lambda^{\nu} Q \quad (11)$$

A konštantné výnosy z rozsahu sú pre  $\nu = 1$ , tie nás budú zaujímať (všeobecne  $\nu > 0$ ).

Parameter distribúcie  $\delta$  určuje rozdelenie produkcie na podiely faktorov. Stačí, ak sledujeme špeciálny prípad  $\nu = 1$  (homogénna CES funkcia), marginálne produkty kapitálu a práce sú:

$$f_K = \frac{\partial Q}{\partial K} = \gamma \left[ \delta K^{-\rho} + (1 - \delta) L^{-\rho} \right]^{-(1+\rho)/\rho} \delta K^{-\rho-1} = \delta \left( \frac{Q}{K} \right)^{1+\rho} \quad (12)$$

$$f_L = \frac{\partial Q}{\partial L} = \gamma \left[ \delta K^{-\rho} + (1 - \delta) L^{-\rho} \right]^{-(1+\rho)/\rho} (1 - \delta) L^{-\rho-1} = (1 - \delta) \left( \frac{Q}{L} \right)^{1+\rho} \quad (13)$$

Podiel týchto dvoch marginálnych produktov sa rovná:

$$\frac{f_L}{f_K} = \frac{(1 - \delta)}{\delta} \left( \frac{K}{L} \right)^{1+\rho} \quad (14)$$

Iný spôsob uvažovania je cez minimalizáciu nákladov:

$$\frac{wL}{rK} = \frac{f_L}{f_K} \frac{L}{K} = \frac{(1-\delta)}{\delta} \left(\frac{K}{L}\right)^\rho \quad (15)$$

Parameter substitúcie  $\rho$  má na starosti stupeň faktora substitúcie. Elasticita substitúcie sa získa z podielu marginálnych produktov, kde platí:

$$\frac{K}{L} = \left(\frac{\delta}{(1-\delta)}\right)^{1/(1+\rho)} \left(\frac{f_L}{f_K}\right)^{1/(1+\rho)} \quad \text{a po zlogaritmovaní: } \sigma = \frac{d \ln(K/L)}{d \ln(f_L/f_K)} = \frac{1}{1+\rho} \quad (16)$$

Pre Leontievovu technológiu sa parameter substitúcie rovná nekonečno a pre Cobba-Douglasa sa rovná 0 a pre dokonalú substitúciu je  $-1$ .

Produkčná funkcia s konštantnou elasticitou substitúcie je inherentne nelineárna – nedá sa transformovať na lineárnu funkciu. Zlogaritmujeme ju a dostaneme výraz:

$$\ln Q = \ln \gamma - \nu/\rho \ln [\delta K^{-\rho} + (1-\delta)L^{-\rho}] \quad (17)$$

ktorý môžeme odhadovať ako nelineárnu funkciu alebo pomocou Taylorovho rozvoja v bode (napr.  $\rho = 0$ ), pričom sa neuvažujú členy tretieho a vyšších rádov:

$$\ln Q = \ln \gamma + \nu\delta \ln K + \nu(1-\delta) \ln L - (1/2)\rho\nu\delta(1-\delta)[\ln K + \ln L]^2 \quad (18)$$

Pri využití metódy maximálnej vierohodnosti nemusíme využívať aproximáciu Taylorovým rozvojom, ale môžeme odhadovať priamo nelineárnu funkciu (17). Po výbere nového objektu typu *LogL* by sme zadali špecifikáciu:

```
@logl LL1
rez = log(Q) - log(c(1)) + (c(2)/c(3)) * log(c(4)*K^(-c(3)) + (1-c(4))*L^(-c(3)))
ll1 = -log(c(5))-(rez^2)/(2*c(5)^2)-0.5*log(2*@acos(-1))
```

Výsledkom je odhad parametra efektívnosti  $\gamma$ , ktorý je zobrazený ako hodnota C(1), odhad stupňa homogenosti  $\nu$  ako hodnota C(2), odhad parametra substitúcie  $\rho$ , ktorý je zobrazený ako hodnota C(3), odhad parametra distribúcie  $\delta$  ako hodnota C(4) a odhad štandardnej odchýlky normálne rozdelených náhodných zložiek, ktorý je zobrazený ako hodnota C(5).

Keďže odhadujeme nelineárny model, kvôli používaným numerickým metódam, je dôležitým nastavenie štartovacích hodnôt parametrov. Preferujeme nastavenie kladných hodnôt obvykle v intervale od 0 po 1.

Hodnotu elasticity substitúcie, ktorá nás z odhadnutých parametrov zaujíma najviac, vypočítame ako  $1/(1+c(3))$ . Zároveň ju môžeme testovať pomocou Waldovho testu (za predpokladu, že *LogL* objekt má názov CES), napríklad či sa nerovná 1:

```
ces.wald 1/(1+c(3))=1
```

## 4 ZÁVER

Metóda maximálnej vierohodnosti, ktorou sa zaoberáme v tomto príspevku, má prominentné postavenie medzi mnohými existujúcimi metódami odhadu parametrov. Viaceré typy modelov sa v špecializovaných ekonometrických programoch, ako je aj EViews, odhadujú práve pomocou tejto metódy. Odhad dvoch najjednoduchších z takýchto modelov, a to model s autokorelovanými náhodnými zložkami a logit model sme prezentovali prostredníctvom nástroja určeného na odhad touto metódou v programe EViews, ktorým je objekt typu *LogL* a bežne používanými zabudovanými prostriedkami v objekte typu *Equation*.

Takýmto spôsobom sme prezentovali možnosti objektu typu *LogL*. Ten sme následne využili na odhad parametrov nelineárneho modelu produkčnej funkcie s konštantnou elasticitou substitúcie, ktorá je predmetom nášho výskumu. Takto získané odhady môžeme porovnávať s výsledkami, ktoré sme dosiahli pri odhade produkčných funkcií inými metódami.

Pri inštalovaní programu EViews sa môže do priečinku s programom zvoliť nainštalovanie aj príkladov z rôznych oblastí, maximálnu vierohodnosť nevynímajúc. V podpriečinku s názvom *Example Files\Sample Programs\logl* nájdeme programy ukazujúce odhad viacerých typov ARCH modelov alebo modelov s kvalitatívnymi závislými premennými ako napríklad probit. Tie môžu rozšíriť nami prezentované modely o náročnejšie oblasti použitia tejto metódy.

### Použitá literatúra

1. GILES, D. 2013. *Maximum Likelihood Estimation in EViews*. Dostupné na: <http://davegiles.blogspot.sk/2013/12/maximum-likelihood-estimation-in-eviews.html> [03.11.2017]
2. GREENE, W. 2011. *Econometric Analysis, 7th ed.* Prentice Hall, 2011.
3. HATRÁK, M. 2007. *Ekonometria*. Bratislava: IURA Edition, 2007.
4. LUKÁČIK, M. 2013. *Parametre produkčnej funkcie ekonomiky*. In: Zborník medzinárodného vedeckého seminára Nové trendy v ekonometrii a operačným výzkumu, Bratislava: Vydavateľstvo EKONÓM, 2013.
5. LUKÁČIK, M., LUKÁČIKOVÁ, A., SZOMOLÁNYI, K. 2013. *Ekonometria 1*. Bratislava: Vydavateľstvo EKONÓM, 2013.
6. LUKÁČIKOVÁ, A. 2011. *Využitie modelu Logit pri modelovaní možnosti poistiť sa*. In: Zborník medzinárodného vedeckého seminára Nové trendy v ekonometrii a operačným výzkumu, Bratislava: Vydavateľstvo EKONÓM, 2011.
7. SZOMOLÁNYI, K. 2014. *Normalizovaná produkčná funkcia s konštantnou elasticitou substitúcie vstupov*. In: Zborník medzinárodného vedeckého seminára Nové trendy v ekonometrii a operačným výzkumu, Bratislava: Vydavateľstvo EKONÓM, 2014.
8. SZOMOLÁNYI, K. 2017. *Elasticita substitúcie vstupov z krátkodobého hľadiska*. In: Zborník zo seminára Využitie kvantitatívnych metód vo vedecko-výskumnej činnosti a v praxi XII, Račkova dolina. Bratislava: Vydavateľstvo EKONÓM, 2017.
9. SZOMOLÁNYI, K., LUKÁČIK, M., LUKÁČIKOVÁ, A. 2017. *Long-run Elasticity of Substitution in the Slovak Economy*. In: Proceedings of the Entrenova '17, ENTerprise REsearch inNOVAtion Conference, Dubrovnik, Croatia. Zagreb: Udruga za promicanje inovacija i istraživanja u ekonomiji "Irenet", pp. 228-232, 2017.
10. User's Guide I, II. – Eviews, IHS Global Inc., Irvine CA.

### Kontaktné údaje

doc. Ing. Martin Lukáčik, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 822

email: [lukacik@euba.sk](mailto:lukacik@euba.sk)



# ODHAD CENY BYTOV VYUŽITÍM EKONOMETRICKÉHO MODELU<sup>1</sup>

## ESTIMATE OF PRICE OF FLATS USING ECONOMETRIC MODEL

*Adriana Lukáčiková*

### **Abstrakt**

V príspevku bude prezentovaný odhad jednorovnicového ekonometrického modelu priemernej ceny bytu, ktorý je vhodný na objasnenie základných princípov pri výučbe úvodného kurzu ekonometrie. Prvý odhadnutý model bude uvažovať s jednou kvantitatívnou vysvetľujúcou premennou, výmerou bytu. Ďalej bude model rozšírený o kvalitatívnu premennú vyjadrujúcu lokalitu bytu, ktorá bude v modeli reprezentovaná umelými premennými.

**KLúčové slová:** *ekonometrický model, umelé premenné, cena bytu*

### **Abstract**

An estimate of single equation econometric model for the estimation of price of flats, which is suitable for clarifying the basic principles at the initial econometrics course are presented in the paper. The first estimated model will consider one quantitative explanatory variable, the size of the flat. Furthermore, the model will be extended by a qualitative variable expressing the location of an apartment, which will be represented by dummy variables in the model.

**Keywords:** *econometric model, dummy variables, price of flat*

## **1 ÚVOD**

Ekonometrické modely, ako tie jednoduchšie, teda jednorovnicové modely, tak aj zložitejšie, viacrovnicové modely sa dajú aplikovať v rôznych oblastiach. Jednak to môže byť potvrdenie alebo vyvrátenie vzťahov odvodených na základe ekonomickej teórie, ale s využitím ekonometrických modelov sa stretávame aj pri riešení problémov v rôznych životných situáciách.

Často sú pri výučbe základného kurzu ekonometrie vhodnejšie nie zložité teoretické ekonomické modely, ale jednoduchšie modely popisujúce reálne rozhodovanie. Jednorovnicový ekonometrický model sa dá aplikovať napríklad aj pri odhade ceny bytov. Inštitút finančnej politiky pri Ministerstve financií Slovenskej republiky publikoval Kalkulačku hodnoty bytov, ktorá je skonštruovaná na základe ekonometrického modelu s veľkým počtom vysvetľujúcich premenných. V rámci výučby ekonometrie by bolo náročné konštruovať model s takým počtom nezávislých premenných, preto budeme uvažovať iba s niektorými vysvetľujúcimi premennými.

---

<sup>1</sup> Príspevok vznikol s podporou projektu VEGA 1/044/15 „Ekonometrická analýza produkčných možností ekonomiky a trhu práce na Slovensku“.

## 2 JEDNOROVNICOVÝ EKONOMETRICKÝ MODEL

### 2.1 Model s jednou kvantitatívnou vysvetľujúcou premennou

V prvom prezentovanom jednorovnicovom modeli budeme uvažovať iba s jednou vysvetľujúcou premennou, a to s výmerou bytu v m<sup>2</sup> (*VYMER*<sub>*i*</sub>). Vysvetľovanou premennou je cena bytu v eurách (*CENA*<sub>*i*</sub>). Model je odhadnutý na základe údajov o 561 bytoch v Slovenskej republike získaných z realitnej kancelárie.

Odhadnutý model má tvar:

$$CENA_{vyp_i} = 12578,5 + 1092,7 * VYMER_{i}$$

Na seminároch je vhodné študentom ukázať interpretáciu parametrov modelu respektíve možnosti použitia modelu. Model verifikujeme štandardným postupom, posudzujeme ekonomickú interpretáciu parametrov modelu, testujeme štatistickú významnosť parametrov modelu, keďže ide o prierezové údaje, testujeme prítomnosť nesplnenia predpokladu o konštantnosti rozptylu náhodnej zložky.

Parameter pri premennej *VYMER* vyjadruje, o koľko eur sa v priemere zvýši cena bytu pri zvýšení výmery o jeden meter štvorcový. Po dosadení do modelu vieme vypočítať, aká bude priemerná cena bytu s danou výmerou napr. 70 metrov štvorcových

$$CENA_{vyp} = 12578,5 + 1092,7 * 70 = 89\ 067,5$$

Podľa tohto modelu by však priemerná cena bytu s výmerou 70 m<sup>2</sup> v Bratislave aj v Prešove resp. Trnave bola rovnaká. Preto do modelu zavedieme kvalitatívnu premennú *KRAJ* vyjadrujúcu lokalitu, v ktorej sa byt nachádza.

Model môžeme modifikovať tak, že ako vysvetľovanú premennú budeme uvažovať cenu za meter štvorcový (*CENAm2*). V takto formulovanom modeli predpokladáme, že parameter pri vysvetľujúcej premennej bude mať záporné znamienko, teda s rastom rozlohy bytu cena za jeden meter štvorcový klesá.

Odhadnutý model má tvar:

$$CENAm2_{vyp_i} = 1532,908 - 3,298 * VYMER_{i}$$

Priemernú cenu bytu s danou výmerou vypočítame tak, že najprv dosadíme do odhadnutého modelu výmeru, čím dostaneme cenu za meter štvorcový a následne dopočítame cenu za celý byt.

### 2.2 Model s kvalitatívnou vysvetľujúcou premennou

Do modelu pridáme kvalitatívnu premennú *KRAJ*. Rozhodnutie uvažovať pri analýze vplyv kvalitatívnych premenných na vysvetľovanú premennú určí, na základe informácie, ktorú majú v modeli obsiahnuť a podľa ich predpokladaného spôsobu vplyvu, akým spôsobom a v akom počte ich pridáme medzi vysvetľujúce premenné umelé premenné.

Keďže uvažujeme s kvalitatívnou premennou, ktorá má osem obmien, do modelu zaradíme sedem umelých premenných. Určíme porovnávajúci kraj, napríklad Bratislavský kraj a pre ostatné kraje vytvoríme umelé premenné. Umelá premenná pre daný kraj bude mať hodnotu 1, ak sa byt v uvažovanom kraji nachádza a hodnotu 0, ak sa byt v kraji nenachádza. Umelé premenné pre každý kraj nazveme skratkou kraja, teda TT- Trnavský kraj, TN- Trenčiansky kraj, NR – Nitriansky kraj, BB – Banskobystrický kraj, ZA – Žilinský kraj, PO – Prešovský kraj a KE – Košický kraj.

Výstup odhadu parametrov modelu z ekonometrického softvéru EViews je uedený na Obrázku 1.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	50838.41	2424.676	20.96709	0.0000
VYMERA	1117.892	22.84872	48.92580	0.0000
TT	-43799.61	2640.924	-16.58496	0.0000
TN	-52684.15	2445.953	-21.53931	0.0000
NR	-39892.06	2591.912	-15.39098	0.0000
ZA	-38122.91	2467.582	-15.44950	0.0000
BB	-45795.64	2492.085	-18.37644	0.0000
PO	-52854.67	2441.047	-21.65246	0.0000
KE	-29576.80	2533.866	-11.67260	0.0000

R-squared	0.843746	Mean dependent var	87236.30
Adjusted R-squared	0.841481	S.D. dependent var	33860.43
S.E. of regression	13481.32	Akaike info criterion	21.87191
Sum squared resid	1.00E+11	Schwarz criterion	21.94137
Log likelihood	-6126.071	Hannan-Quinn criter.	21.89903
F-statistic	372.5886	Durbin-Watson stat	1.363406
Prob(F-statistic)	0.000000		

Obrázok 1 Výstup z EViewsu

Odhadnutý model má tvar:

$$CENA_{vyp_i} = 50838,4 + 1117,9 * VYMERA_i - 43799 * TT - 52684 * TN - 39892 * NR - 38123 * ZA - 45795 * BB - 29577 * KE - 52854 * PO$$

Parameter pri vysvetľujúcej premennej má analogickú interpretáciu ako v prvom odhadovanom modeli. Odhadnuté parametre pri umelých premenných majú ako sme mohli predpokladať už pred odhadom záporné znamienka, čo vyjadruje, že byty v ostatných krajoch sú v porovnaní s Bratislavským krajom lacnejšie. Po Bratislavskom kraji nasleduje Košický kraj. Najlacnejšie byty sú v Prešovskom kraji, potom nasleduje Trenčiansky kraj.

Na základe odhadnutého modelu vieme odpovedať na otázky typu: Ako vypočítame priemernú cenu bytu s danou výmerou, napr. 100m<sup>2</sup> v Bratislavskom kraji? Ako vypočítame priemernú cenu bytu s danou výmerou, napr. 80 m<sup>2</sup> v Prešovskom kraji?

### 3 ZÁVER

Prezentovaný model je príkladom, ako možno na seminároch úvodného kurzu ekonometrie ukázať, akým spôsobom sa formulujú ekonometrické modely, v ktorých vystupujú ako vysvetľujúce premenné kvantitatívne premenné a kvalitatívne premenné, v akom tvare sú tieto premenné reprezentované v modeli. Dôležité je tiež poukázať na rozdiel v interpretácii parametrov pri kvantitatívnych a kvalitatívnych respektíve umelých premenných.

Uvedený model by bolo vhodné rozšíriť o ďalšie vysvetľujúce premenné, a to napríklad stav bytu, vek bytu, či sa byt nachádza na prízemí alebo sa jedná o byt na najvyššom poschodí.

Lokalita bytu sa dá uviesť na úroveň okresov alebo diferencovať, či ide o byt v sídle nad určitý počet obyvateľov alebo nie. Napriek zjednodušeniu v uvádzanom modeli, je tento model vhodný ako súčasť výučby ekonometrie na prvom stupni štúdia.

### **Použitá literatúra**

1. LUKÁČIKOVÁ, A. – LUKÁČIK, M. 2008: *Ekonometrické modelovanie s aplikáciami*. Bratislava: Ekonóm 2008.
2. LUKÁČIKOVÁ, A. – LUKÁČIK, M – SZOMOLÁNYI, K. 2013: *Ekonometria I*. Bratislava: Ekonóm 2013.
3. WOOLDRIDGE, J. M. 2009: *Introductory Econometrics. A Modern Approach*. Mason, USA: South – Western Cengage Learning 2009.
4. [http://www-ext.nbs.sk/\\_img/Documents/PUBLIK/MU/Car09-03-Biatec.pdf](http://www-ext.nbs.sk/_img/Documents/PUBLIK/MU/Car09-03-Biatec.pdf), 11. 10. 2017
5. <http://www.institutfinancnejpolitiky.sk/kalkulacky/nehnutelnosti>, 11. 10. 2017

### **Kontaktné údaje**

Ing. Adriana Lukáčiková, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 824

email: [adriana.lukacikova@euba.sk](mailto:adriana.lukacikova@euba.sk)

# PROSTOROVÁ EKONOMETRIE V MAKROEKONOMICKÉ ANALÝZE

## SPATIAL ECONOMETRICS IN MACROECONOMIC ANALYSIS

Simona Macková

### Abstrakt

Prostorová ekonometrie může přinést užitečný přístup v makroekonomické analýze regionálních dat. V příspěvku jsou nastíněny vhodné modely průřezových dat zohledňující jejich geografickou polohu. Pro analýzu je využita relace sousednosti reprezentovaná maticí prostorových vah. Základní modely aplikujeme na reálná makroekonomická data a prověřujeme, jak popisují vztah mezi příjmy domácností, HDP a mírou nezaměstnanosti v západní Evropě. Výsledky jsou porovnány s lineárním regresním modelem.

**Klíčová slova:** prostorové modely, makroekonomická analýza

### Abstract

Spatial econometrics can bring a useful approach in macroeconomic analysis of regional data. In this paper suitable cross-section data models regarding their geographical location are delineated. Neighbourhood relation is used for the analysis. The relation of neighbourhood among the regions is represented by spatial weight matrix. Basic models are applied on real macroeconomic data and relations between household incomes, GDP and unemployment rate in western Europe are tested. Results are compared with a linear regression model.

**Key words:** spatial models, macroeconomic analysis

### Geo-Data source and copyright acknowledgement

GISCO NUTS is a geographical dataset developed by the European Commission based on EuroBoundary Map (EBM) from EuroGeographics.

When the GISCO NUTS geographical dataset is used in any printed or electronic publication, the source data set shall be acknowledged in the legend of the map and in the introductory page of the publication.

The data source notice on the legend of the map and the introductory part of the publication is as follows:

Data source: GISCO - Eurostat (European Commission)

Administrative boundaries: © EuroGeographics.

# Úvod

Prostorová ekonometrie se zabývá modely, které mohou být vhodné pro analýzu prostorových dat zejména makroekonomických ukazatelů. Mezi pozorovanými jednotkami definujeme vztah sousednosti, který vyjadřujeme maticí prostorových vah. Testujeme či vyvracíme vliv geografické polohy sledovaných jednotek a zadefinujeme některé prostorové modely.

Praktická aplikace je věnována oblasti makroekonomické analýze západní Evropy.

## 1 Prostorové modely

### 1.1 Matice vah

Matice vah  $\mathbf{W} = w_{ij}$  pro  $i, j = 1, 2, \dots, n$  je důležitým aspektem prostorové ekonometrie. Měla by být sestavována s ohledem na sledované prostředí, neboť může zásadně ovlivnit výsledek analýzy. Matici  $\mathbf{W}$  nemůžeme odhadovat a musí být sestavena ještě před dalšími kroky.

V tomto příspěvku matici vah sestavíme na základě metody  $k$  nejbližších sousedů. Jako centroidy jednotlivých regionů volíme geografické středy. Řádkovou normalizací matice sousednosti, která je maticí nul a jedniček, dostáváme matici  $\mathbf{W}$ . Tu následně využijeme v prostorových modelech pro zohlednění relace sousednosti.

Pro další postup je třeba prověřit prostorovou autokorelaci například pomocí Gearyho testu, či Moranova indexu. O testování prostorové autokorelace blíže pojednává například článek LeSage a Pace (2009).

### 1.2 Vybrané prostorové modely

#### Model GNS

Plný prostorový model pro průřezová data, neboli GNS model, zohledňuje všechny interakční efekty a mohl by být popsán následovně

$$\begin{aligned} \mathbf{Y} &= \delta \mathbf{W} \mathbf{Y} + \alpha \mathbf{1}_N + \mathbf{X} \boldsymbol{\beta} + \mathbf{W} \mathbf{X} \boldsymbol{\theta} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\epsilon}. \end{aligned} \tag{1}$$

Matice  $\mathbf{W}$  označuje již zmiňovanou nezápornou matici prostorových vah ( $N \times N$ ), která popisuje vlivy regionů mezi sebou. Dále uvažujeme tři typy interakcí, které zohlední, jak mohou pozorování záviset na pozorování v ostatních oblastech.  $\mathbf{W} \mathbf{Y}$  označuje endogenní interakční efekt mezi závislou proměnnou. Například, míra nezaměstnanosti v  $i$ -té jednotky je závislá na míře nezaměstnanosti ostatních jednotek.  $\mathbf{W} \mathbf{X}$  je exogenní interakční efekt mezi nezávisle proměnnou. Příkladem uveďme závislost míry nezaměstnanosti v  $i$ -té oblasti na příjmech domácností v ostatních regionech. Nakonec  $\mathbf{W} \mathbf{u}$  popisuje interakční efekt mezi náhodnou složkou.

Na základě prostorových modelů budeme odhadovat  $\alpha$ ,  $\delta$ ,  $\lambda$  a vektory parametrů  $\beta$  a  $\theta$ . Jednou z obvyklých metod pro odhad parametrů je metoda maximální věrohodnosti. Aplikaci na prostorové modely je věnována pasáž v knize LaSage a Pace (2009).

U modelu GNS se často setkáváme s přeparametrizováním. Proto je vhodné některé interakce vynechat. Podíváme na nejčastěji využívané modely.

## Model SAR

Tento model je též nazýván prostorový autoregresní proces. Součtem  $\sum_{j=1}^N w_{ij}Y_j$  označujeme tak zvané prostorové zpoždění. Koeficient  $\delta$  vyjadřuje sílu prostorové závislosti ve sledovaném vzorku. Modely SAR vznikají z plného prostorového modelu tak, že parametry  $\lambda$  a  $\theta$  položíme rovné nule. To znamená, že neuvažujeme prostorovou závislost nezávisle proměnné ani náhodné složky.

$$\mathbf{Y} = \delta \mathbf{W} \mathbf{Y} + \alpha \mathbf{1}_N + \mathbf{X} \beta + \epsilon \quad (2)$$

$\alpha$  reprezentuje absolutní člen rovnice,  $\epsilon$  je vektor náhodné složky, respektive bílý šum. Pokud by i parametr  $\delta$  byl roven nule, tak se dostáváme ke klasické lineární regresi bez prostorové závislosti.

Model SAR zachycuje situaci, kdy vektor  $\mathbf{Y}$  reprezentující příjmy domácností v jednotlivých regionech závisí na příjmech v sousedních oblastech a také na míře nezaměstnanosti. Není však již závislý na míře nezaměstnanosti u sousedů.

## Model SEM

Dalším využívaným modelem je *spatial error model*, SEM. Zabýváme se prostorovou závislostí náhodné složky nikoliv pozorované proměnné. Chyba  $\mathbf{u}$  se řídí modelem SAR, je to analogie modelu (2). Volba takového modelu naznačuje, že existuje další proměnná, která není v modelu zahrnuta, ale je prostorově závislá. Tento model dostáváme, pokud v modelu (1) položíme koeficienty  $\delta$  a  $\theta$  nule. Zůstává  $\lambda$  jako prostorový parametr.

$$\begin{aligned} \mathbf{Y} &= \alpha \mathbf{1}_N + \mathbf{X} \beta + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W} \mathbf{u} + \epsilon \end{aligned} \quad (3)$$

### 1.3 Hodnotící kritéria

Při vyhodnocování výsledků samozřejmě požadujeme, aby odhady koeficientů, případně přímých a nepřímých efektů byly statisticky významné. Neměla by být přítomna heteroskedasticita a prostorová autokorelace reziduí. Pro měřitelné porovnání modelů využijeme logaritmus věrohodnostní funkce, Akaikeho informační kritérium a Bayesovo informační kritérium. Popořadě značeno *LogLik*, *AIC* a *BIC*.

## 2 Makroekonomická analýza

Analýza je zaměřena na státy západní Evropy, a to Německo, Francie, Nizozemsko, Belgie a Lucembursko. Přestože jsou si tyto země geograficky velice blízké, jejich ekonomické situace jsou rozdílné. Významné rozdíly pozorujeme nejen napříč státy, ale i v regionech jednotlivých zemí. Sledujeme celkem 83 regionů. Následná analýza dat bude provedena v softwaru R. Jak provádět základní vyhodnocení prostorové závislosti je popsáno a na příkladech ukázáno v článku z university v Illinois Anselin (2003).

Vycházíme z datových souborů z webové databáze Eurostatu, statistického úřadu Evropské unie z roku 2011.

HDP – Využíváme výši HDP v eurech na osobu.

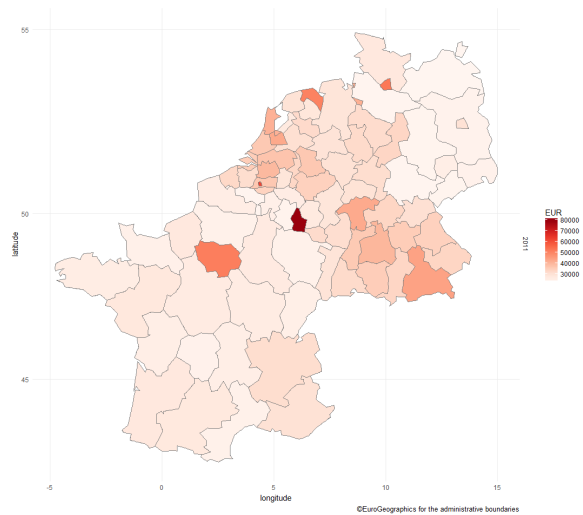
Nezaměstnanost – Popisuje procento nezaměstnaných z populace osob v produktivním věku.

Příjmy domácnosti – Jedná se o primární příjem domácností v eurech na osobu.

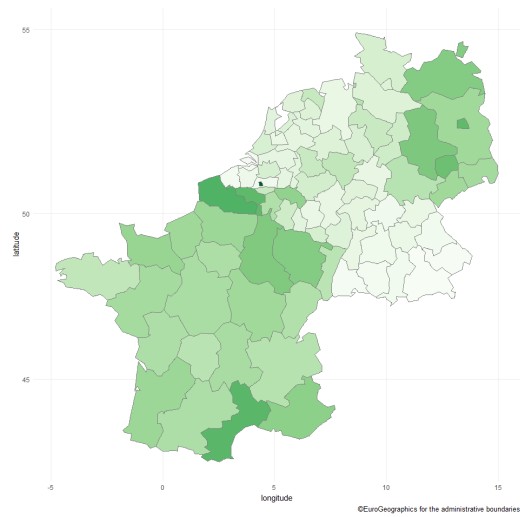
Podle obrázků 1, 2, 3 můžeme zhodnotit, jak si stojí jednotlivé regiony v ekonomické situaci.

Ve všech měřítkách vyčnívá Lucembursko, které má dlouhodobě nejvyšší HDP na osobu na světě, velice nízkou míru nezaměstnanosti a vysoké příjmy domácností. Na území Německa pozorujeme stále dopad dřívějšího dělení na Spolkovou republiku Německo a Německou demokratickou republiku. Bývalá NDR stále vykazuje nižší HDP i příjmy domácností a vyšší míru nezaměstnanosti. Na druhé straně Francie se jeví na celém svém území homogenně s výjimkou prosperujícího okolí Paříže.

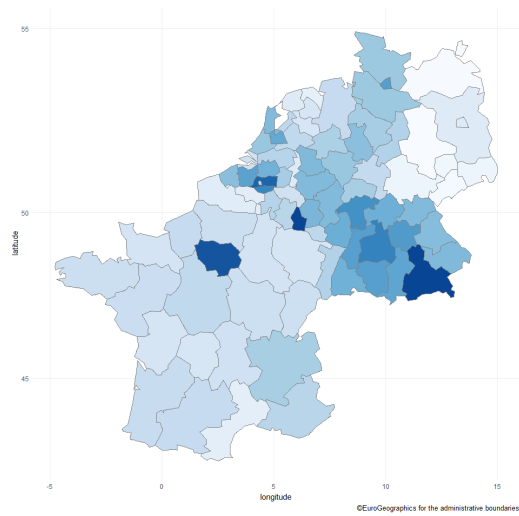




Obrázek 1: HDP na osobu v EUR v západní Evropě



Obrázek 2: Míra nezaměstnanosti v % v západní Evropě



Obrázek 3: Příjem domácností na osobu v EUR v západní Evropě

## 2.1 Porovnání modelů

Jak již bylo nastíněno, zaměříme se na dva prostorové modely SAR a SEM. Jejich výsledky porovnáme s klasickým lineárním regresním modelem, který žádnou prostorovou závislost nezohledňuje. Zadefinujeme si modely, jak budou vypadat pro naši analýzu.

### Lineární regresní model

$$INC = \alpha + \beta_1 GDP + \beta_2 UNE + \epsilon \quad (4)$$

### Model SEM

$$INC = \alpha + \beta_1 GDP + \beta_2 UNE + \mathbf{u} \quad (5)$$

$$\mathbf{u} = \lambda \mathbf{W} \mathbf{u} + \epsilon$$

Mimo jiné budeme v tomto případě testovat statistickou významnost koeficientu  $\lambda$ . Bude-li signifikantní, má význam do modelu zahrnout součin  $\mathbf{W} \mathbf{u}$ , který popisuje prostorovou závislost. Typ modelu SEM naznačuje, že v náhodné složce je nějaký další činitel, který závislost vysvětluje.

### Model SAR

Opět budeme v modelu uvažovat HDP i míru nezaměstnanosti a následně SAR model můžeme psát

$$INC = \alpha + \delta \mathbf{W} INC + \beta_1 GDP + \beta_2 UNE + \epsilon \quad (6)$$

Díky testům prostorové autokorelace můžeme potvrdit její přítomnost, a tedy se vrhnout na další kroky analýzy. Pro nalezení vhodného modelu lze využít test Lagrangeových multiplikátorů. Ten se zaměřuje na srovnání modelů SEM a SAR. Princip tohoto testu je detailně vysvětlen v článku Florax a Nijkamp (2003). Nám tento test doporučil model SEM. Přesto vyzkoušíme oba výše zmiňované modely a výsledky srovnáme.

Tabulka 1 shrnuje výsledky jednotlivých modelů.

Odhadnuté hodnoty regresních koeficientů jsou v prvních řádcích. Dále uvádíme hodnoty Akaikeho informačního kritéria a logaritmus věrohodnostní funkce a Moranův test prostorové nezávislosti reziduí v odhadnutém modelu. Dvě hvězdičky u koeficientů značí statistickou významnost na hladině 5%, jedna hvězdička na hladině 10%. Koeficienty bez hvězdičky nepovažujeme za statisticky významné. Dvě kolečka znamenají, že na hladině 5% nezamítáme nulovou hypotézu  $H_0$ , jedno kolečko na hladině 1%.

Odhady koeficientů vycházejí v souladu s naším očekáváním. S rostoucím hrubým domácím produktem v regionu rostou i příjmy domácností a s rostoucí mírou nezaměstnanosti klesají.

V obou případech prostorových modelů nezamítáme hypotézu o prostorové nezávislosti reziduí. Hodnotící kritéria u modelu SEM vycházejí nejlépe: *AIC* je nejnižší a hodnota *LogLik* je nejvyšší. Dle tohoto hodnocení bychom jako nejlepší zvolili model SEM. Nicméně ještě se podíváme na významnost přímých a nepřímých efektů modelu SAR, které jsou důležité pro vyhodnocení modelů tohoto typu. Teorie přímých a nepřímých efektů je podrobněji zpracována v článku Elhorst (2014).

	OLS	SEM	SAR
$\alpha$	$2.02 \cdot 10^{4**}$	$2.15 \cdot 10^{4**}$	$1.38 \cdot 10^{4**}$
$\beta_1$	0.1881**	0.1748**	1.7784**
$\beta_2$	-597.40**	-732.10**	-497.47**
$\delta$			0.2754**
$\lambda$		0.5496**	
AIC	1512.1	1495.2	1506.5
LogLik	-752.04	-742.62	-748.35
Moran		-0.0362 <sup>oo</sup>	0.1396 <sup>oo</sup>

\*\* Významné na hladině 5%, \* Významné na hladině 10%

<sup>oo</sup> Nezamítáme  $H_0$  na hladině 5%, <sup>o</sup> Nezamítáme  $H_0$  na hladině 1%

Tabulka 1: Porovnání modelů vysvětlujících příjmy domácností

Pro model SAR jsou rozdíly mezi odhady přímých efektů a odhady koeficientů poměrně malé. U klasického lineárního modelu a modelu SEM jsou úplně stejné. Pro model autoregresní prostorový se efekty liší v důsledku endogenního interakčního efektu  $WY$ . Tato interakce způsobuje tak zvaný zpětný efekt, *feedback effect*. Můžeme jej interpretovat, že změna výše HDP ve vybraném regionu působí na výši HDP u jeho sousedů a ta se následně projeví i ve formě zpětného efektu na region, ve kterém byla změna podnícena. Přímé efekty odhadnuté různými modely jsou řádově stejné. Nabývají hodnot od -732,10 po -508,42 pro výši HDP. Máme pouze jeden model, kde vyčíslujeme nepřímé efekty. Nemáme tedy pro ně srovnání. Oba odhady nepřímých efektů modelu SAR vyšly statisticky významné na hladině 5%.

	OLS	SEM	SAR
<i>Přímý efekt</i>			
GDP	0.1881**	0.1748**	0.1818**
UNE	-597.40**	-732.10**	-508.47**
<i>Nepřímý efekt</i>			
GDP			0.0637*
UNE			-178.07**

\*\* Významné na hladině 5%, \* Významné na hladině 10%

Tabulka 2: Porovnání přímých a nepřímých efektů

Závěrem bychom mohli zhodnotit, že dle vyhodnocení na základě informačních kritérií se test Lagrangeových multiplikátorů nemýlil a jako lepší se jeví model SEM. Nicméně se ukázal statisticky významný i nepřímý efekt a ten není tímto modelem sledován. V tuto chvíli by bylo ideální věnovat se hlubší analýze a otestovat i další možné prostorové modely jako je například Durbinův model.

## Závěr

Úvodní část příspěvku byla věnována problematice matice prostorových vah  $W$  a základním modelům prostorové ekonometrie. Z celé škály prostorových modelů jsme věnovali pozornost modelům SEM, SAR, které zohledňují pouze některé prostorové interakce.

Je třeba si uvědomit, že odhady prostorových modelů nemůžeme vždy interpretovat stejně jako u klasického lineárního modelu. Odhadujeme pro každou pozorovanou jednotku zvlášť přímé a nepřímé efekty, které popisují vliv v rámci sledované oblasti i mezi oblastmi navzájem.

V ukázkovém příkladě makroekonomické analýzy jsme porovnali výsledky modelů SEM, a SAR. Z hlediska hodnotících kritérií vždy došlo ke zlepšení ve srovnání s výchozím modelem lineární regrese. Je třeba zvážit, zda toto zlepšení je pro naši požadovanou analýzu zásadní a vyplatí se nám prostorové modely využívat.

## Použitá literatura

ANSELIN, L. (2003). An Introduction to Spatial Regression Analysis in R. *Urbana-Champaign*

ELHORST, J.P. (2014) *Spatial Econometrics: from cross-section data to spatial panels*. Springer, Heidelberg. ISBN 978-80-200-1707-9.

EUROSTAT. European Commission: Your key to European statistics.  
URL <http://ec.europa.eu/eurostat/data/database>.

FLORAX, R.J.G.M. a NIJKAP, P. (2003) Misspecification in Linear Spatial Regression Models. *Tinbergen Institute Discussion Paper*, (TI 2003-081/3)

LESAGE, J. a PACE, R.K. (2009) *Introduction to Spatial Econometrics*. CRC Press, U.S.A. ISBN 978-1-4200-6424-7.

## Kontaktní údaje

Ing. Simona Macková  
Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky  
Náměstí Winstona Churchilla 4, 130 67 Praha 3  
e-mail: [simona.mackova@vse.cz](mailto:simona.mackova@vse.cz)

# CENOVÉ MODELY REVENUE MANAGEMENTU

## PRICE MODELS IN REVENUE MANAGEMENT

*Renata Majovská, Petr Fiala*

### **Abstrakt**

Revenue management je umění a věda o předvídání zákaznické poptávky v reálném čase a o optimalizaci cen a dostupnosti produktů v závislosti na poptávce. Tento příspěvek je věnován cenovým modelům v revenue managementu. Cenotvorba v revenue managementu je složitý problém. Některé přístupy řeší problém v zjednodušené formě. V příspěvku jsou navrženy deterministické a stochastické modely. Pro stochastické modely je možno používat statické i dynamické přístupy.

***Klíčová slova:** revenue management, cena, deterministické modely, stochastické modely*

### **Abstract**

Revenue management (RM) is the art and science of predicting real-time customer demand and optimizing the price and availability of products according to the demand. This paper is dedicated to price models in revenue management. Pricing in revenue management is a complex problem. Some approaches address the problem in a simplified form. The paper proposes deterministic and stochastic models. For stochastic models, both static and dynamic approaches can be used.

***Keywords:** revenue management, price, deterministic models, stochastic models*

## **1 ÚVOD**

Revenue management se zabývá rozhodnutími, která se týkají řízení poptávky, metodologií a systémy, které tato rozhodnutí umožňují, s cílem dosáhnout zvýšení příjmů (viz např. Talluri a van Ryzin, 2004, Phillips, 2005, Fiala, 2015). Revenue management se týká tří základních kategorií rozhodnutí v oblasti managementu poptávky:

- cenová rozhodnutí,
- rozhodnutí o množství,
- strukturální rozhodnutí.

Cenová rozhodnutí se věnují stanovení cen, jak stanovit ceny v rámci kategorií produktu, během času, jak stanovit ceny v průběhu životního cyklu produktu atd. Rozhodnutí o množství se zabývají otázkami přijetí nebo odmítnutí kupní nabídky, alokace výstupu nebo kapacit jednotlivých segmentů, produktů a prodejních kanálů, stažení produktu z trhu s možností pozdějšího prodeje atd. Strukturální rozhodnutí se týkají výběru prodejních forem, použitých segmentačních a diferenciačních mechanismů, podmínek prodeje, vytváření „balíčků“ produktů atd.

V příspěvku jsou navrženy deterministické a stochastické modely pro tvorbu cen. Pro stochastické modely je možno používat statické i dynamické přístupy. Navržené postupy jsou zjednodušující a je možné je používat při určitých předpokladech a je možné je používat jako heuristické přístupy.

## 2 OBECNÝ MODEL

Obecný model revenue managementu (Bitran, Caldentey, 2003) poskytuje celkový pohled na různé prvky a jejich vzájemné vztahy:

- dodávka,
- produkt,
- informace,
- poptávka.

Prodávající má pevnou částku počáteční kapacity, která se používá k uspokojení cenově závislé poptávky během určitého období prodeje  $[0, T]$ . Tato počáteční kapacita je modelována prostřednictvím  $m$ -rozměrného vektoru  $m$  zdrojů. Kapacitu lze interpretovat například jako volná místa pro určitý let (start-cíl) v daném dni, volné pokoje v hotelu apod. Kapacita je v podstatě daná a prodávající je výlučně určen k nalezení nejlepšího způsobu, jak ji prodat. Z cenového hlediska existují dva důležité atributy dostupné kapacity, její flexibilita a její pomíjivost. Flexibilita měří schopnost produkovat a nabízet různé produkty s využitím původní kapacity  $C_0$ . Pomíjivost se vztahuje k neschopnosti zachovat kapacity v průběhu času. Jak čas postupuje a zdroje jsou spotřebovávány, kapacita se snižuje. Dostupnou kapacitu v čase  $t$  označíme  $C_t = (c_1(t), \dots, c_m(t))$ .

Produkt je podmnožina dostupných zdrojů. Matice  $\mathbf{A} = [a_{ij}]$  je matice rozměru  $(m, n)$  a definována tak, že  $a_{ij}$  představuje množství zdroje  $i$  použitého k produkci jedné jednotky produktu  $j$ . Každý sloupec  $j$  matice  $\mathbf{A}$  představuje jiný produkt a množina  $M = \{A_1, \dots, A_n\}$  je nabídka produktů, nabízených prodávajícím.

Znalost systému a jeho vývoje v čase jsou velmi důležité pro všechny dynamické cenové politiky. Zkoumaná historie  $H_t$  prodejních procesů je definována jako soubor všech relevantních informací dostupných v čase  $t$ , vzhledem k počáteční kapacitě  $C_0$ , nabídce produktů  $M$  a poptávkovým a cenovým procesům. Tato historie by měla zahrnovat alespoň zkoumaný proces poptávky a dostupnou kapacitu, což může zahrnovat i některé další informace, jako například prognózy poptávky.

Množina potenciálních zákazníků je rozdělena do různých segmentů, každý z nich má svou vlastní množinu atributů. Je definován  $d$ -rozměrný stochastický proces  $\mathbf{N}(t, H_t) = (N_1(t, H_t), \dots, N_d(t, H_t))$  kde  $N_j(t, H_t)$  je kumulovaná potenciální poptávka až do času  $t$  segmentu  $j$  vzhledem k dostupným informacím  $H_t$ . Matice  $\mathbf{B}(P) = [b_{ij}]$  je matice rozměru  $(n, d)$  a definována tak, že  $b_{ij}$  představuje jednotky produktu  $i \in M$  požadované zákazníkem v segmentu  $j = 1, \dots, d$ . Poptávka závisí na cenové politice  $P = \{p_t, t \in [0, T]\}$ , kde  $p_t(i, H_t)$  je cena produktu  $i \in M$  v čase  $t$  daná historií  $H_t$ . Efektivní kumulativní poptávkový proces v periodě  $[0, t]$  na produktové úrovni je definován jako  $n$ -rozměrný vektor  $\mathbf{D}(t, P, H) = \mathbf{B}(P)\mathbf{N}(t, H_t)$ . Množina všech přípustných cenových politik, které splňují všechna příslušná omezení, je označena jako  $\Pi$ . Prodávající má schopnost částečně uspokojit poptávku, pokud je to pro něj ziskové. Je definován  $n$ -rozměrný vektor  $\mathbf{S}(t)$ , který představuje kumulativní prodej až do času  $t$ .

Obecný problém revenue managementu je najít řešení pro následující úlohu optimálního řízení

$$\sup_{P, S} E_N \left[ \int_0^T p_t dS(t) \right]$$

při omezeních

$$C_t = C_0 - \mathbf{AS}(t) \geq 0 \text{ pro všechna } t \in [0, T],$$

$$0 \leq \mathbf{S}(t) \leq \mathbf{D}(t, P, H_t) \text{ pro všechna } t \in [0, T],$$

$$P \in \Pi, \text{ and } \mathbf{S}(t) \in H_t.$$

### 3 DETERMINISTICKÉ MODELY

Deterministické modely předpokládají, že prodávající má dokonalé informace o procesu poptávky. Deterministické modely se snadno analyzují a poskytují dobrou aproximaci pro realističtější avšak komplikovanější stochastické modely. Deterministická řešení jsou v některých případech asymptoticky optimální pro daný problém se stochastickou poptávkou (Cooper, 2002).

Nejjednodušší deterministický model uvažuje případ monopolního prodeje jednoho produktu s cenově závislou poptávkou během období  $[0, T]$ . Počáteční zásoba je  $C_0$ , poptávka je deterministická časově závislá s cenově závislou intenzitou  $\mu(p, t)$ . U funkce revenue intenzity  $r(p, t) = p\mu(p, t)$  se předpokládá, že je konkávní stejně jako ve většině reálných situací.

Obecný problém revenue managementu může být zapsán v tomto případě následovně:

$$\max_p \int_0^T p_t \mu(p_t, t) dt$$

při omezení

$$\int_0^T \mu(p_t, t) dt \leq C_0.$$

Toto je standardní problém variačního počtu. Podmínka optimality je dána následovně:

$$p_t^* = \lambda - \frac{\mu(p_t^*, t)}{\mu_p(p_t^*, t)},$$

kde  $\lambda$  je Lagrangeův multiplikátor pro omezení,  $\mu_p$  je parciální derivace  $\mu$  podle ceny  $p$ .

V případě časově homogenní intenzity poptávky ( $\mu(p, t) = \mu(p)$ ), může být dokázáno, že existuje pevná cena, která je optimální po celé období prodeje  $[0, T]$ .

Nechť  $p^* = \operatorname{argmax} \{p\mu(p) : p \geq 0\}$  je cenová politika, která maximalizuje revenue intenzitu a  $\mu^* = \mu(p^*)$  je odpovídající intenzita poptávky. Stejně tak, ať  $p^0$  je řešením pro  $\mu(p^0) T = C_0$  a  $\mu^0 = \mu(p^0)$   $p^0$  je odpovídající intenzita poptávky.

Problém revenue managementu s jediným produktem, s homogenní intenzitou poptávky  $\mu(p)$ , konkávní revenue intenzitou  $r(p) = p\mu(p)$ , má řešení:

- jestliže  $\mu^* T \leq C_0$ , pak optimální cena je  $p^*$  a optimální revenue se rovná  $p^* \mu^* T$ .
- jestliže  $\mu^* T > C_0$ , pak optimální cena je  $p^0$  a optimální revenue se rovná  $p^0 C_0$ .

Tento výsledek je používán jako stavební blok pro konstrukci heuristik a mezí pro stochastické modely.

## 4 STOCHASTICKÉ MODELY

Cenové politiky se stochastickou poptávkou jsou složitější a obtížnější pro výpočet než s deterministickou poptávkou. Na druhou stranu stochastické modely jsou jasně vhodnější k popisu reálných situací, kdy poptávka a zásoby jsou nepředvídatelné v průběhu času a manažeři jsou nuceni reagovat dynamicky přizpůsobováním cen podle toho, jak se projevuje nejistota. Přírozený způsob, jak řešit problémy tohoto typu, je pomocí technik stochastického dynamického programování. V každém rozhodovacím bodě během prodejní sezóny manažer shromažďuje všechny relevantní informace o aktuálním stavu zásob a prodeje a stanoví ceny, za které by produkty měly být prodány. Většina výzkumu byla orientována na případy s jediným produktem při markovských předpokladech o procesu poptávky. V tomto případě jsou relevantní informace pouze úrovně zásob, které manažeři potřebují pro rozhodnutí o cenách.

V případě jediného produktu počáteční kapacita  $C_0$  je skalární veličina, představující počet jednotek produktu, které jsou k dispozici v čase  $t = 0$ . Hodnotová funkce  $V_t(C_t)$  je definována v čase  $t$ , jestliže je zásoba  $C_t$ , potom  $V_t(C_t)$  je optimální očekávaná hodnota příjmu od času  $t$  do konce periody za předpokladu, že aktuální výše zásob v čase  $t$  je  $C_t$ . Čas  $t$  je modelován v literatuře buď jako spojitá nebo diskrétní proměnná. Z praktického hlediska, budou manažeři revidovat svá cenová rozhodnutí pouze v diskrétních časových bodech. Avšak bouřlivý růst Internetu a e-obchodu umožňují použití spojitého modelu pro praktické použití.

### *Statický přístup*

Nejjednodušší přístup k problému je řešení se statickou cenou. Cenová politika je omezena pevnou cenou po celou periodu, tj.  $p_t = p$  pro všechna  $t \in [0, T]$ . Tento typ statické politiky je vhodný pro produkty, které mají jednu nebo více z následujících vlastností:

- krátké prodejní období,
- vysoké náklady na změny cen,
- právní předpisy, které stanovují pevné ceny.

Model pevné ceny je jednoduchý a je možno jej snadno implementovat a řídit. Proto, i když jsou možné změny cen, manažeři se často rozhodnou používat přístup se statickou pevnou cenou. Model pevné ceny je v některých situacích asymptoticky optimální. V tomto modelu s jediným produktem a pevnou cenou je dána hodnotová funkce

$$V(C_0, T) = \max_{p \geq 0} V(C_0, p, T) = \max_{p \geq 0} E[p \min \{D(p, T); C_0\}],$$

kde  $D(p, T)$  je náhodná veličina, představující kumulativní poptávku v periodě  $[0, T]$  při pevné ceně  $p$ . Řešení tohoto problému v uzavřené formě není k dispozici pro obecný případ libovolného rozdělení  $D(p, T)$ . Optimální cenu lze však charakterizovat z hlediska pružnosti poptávky. Necht'  $f(D, p, T)$  je hustota pravděpodobnosti poptávky  $D(p, T)$ . Pružnost poptávky vzhledem k ceně je definována jako

$$e = \frac{pf_p(D, p, T)}{f(D, p, T)},$$

kde  $f_p(D, p, T)$  je parciální derivace  $f(D, p, T)$  podle ceny  $p$ .

Podmínka prvního řádu pro optimalitu řešení je dána vztahem

$$\frac{E[\min\{D; C_0\}e(D, p, T)]}{E[\min\{D; C_0\}]} = -1.$$



Vážená očekávaná hodnota pružnosti musí být rovna -1, kde váha je dána úrovní prodeje  $\min\{D; C_0\}$ .

### **Dynamický přístup**

Model s dynamickou cenou je spojitý časový model, kdy poptávka je popsána Poissonovým procesem s pevnou intenzitou  $\mu$ . Příchozí zákazník v čase  $t$  má pro produkt rezervační cenu  $r_t$ . Z pohledu prodejce je rezervační cena  $r_t$  náhodná veličina s rozdělením  $F(r, t)$ . Je možno uvažovat dva případy (Kincaid a Darling, 1963). V prvním případě prodávající nestanoví ceny, ale přijímá nabídky od potenciálních příchozích kupců, které buď přijme, nebo odmítne. Předpokládá se, že příchozí zákazníci nabídnou svoji rezervační cenu  $r_t$ . V druhém případě, prodávající stanoví cenu  $p_t$  a příchozí zákazníci zakoupí produkt pouze tehdy, jestliže  $p_t \leq r_t$ . Poptávkový proces je v tomto případě Poissonův proces s intenzitou  $\mu(1 - F(p_t, t))$ . Jsou odvozeny podmínky optimality pro hodnotovou funkci  $V_t(C_t)$  a optimální cena  $p_t(C_t)$  pro oba případy

$$p_t(C_t) = \frac{1 - F(p_t(C_t), t)}{f(p_t(C_t), t)} + V_t(C_t) - V_t(C_t - 1).$$

Problém výpočtu optimální cenové strategie se redukuje na výpočet nákladů ušlé příležitosti  $V_t(C_t) - V_t(C_t - 1)$ . Obecně neexistuje přesná uzavřená forma řešení pro optimální cenovou strategii. Může být dokázáno, že

$$\frac{V_t^{\text{det}}(C_t)}{V_t(C_t)} \geq 1 - \frac{1}{2\sqrt{\mu(p^{\text{det}})(T-t)}},$$

kde  $\mu(p^{\text{det}})$  je intenzita poptávky při ceně  $p^{\text{det}}$ .

## **5 ZÁVĚR**

Navržené postupy jsou zjednodušující a je možné je používat při určitých předpokladech a je možné je používat jako heuristické přístupy. Deterministické modely předpokládají, že prodávající má dokonalé informace o procesu poptávky. Deterministické modely je možné snadno analyzovat a poskytují dobrou aproximaci pro realističtější, avšak komplikovanější stochastické modely. Stochastické modely jsou vhodnější k popisu reálných situací. Cenové politiky se stochastickou poptávkou jsou složitější a obtížnější pro výpočet než politiky s deterministickou poptávkou. Lze používat statické nebo dynamické přístupy. Heuristika deterministické ceny je asymptoticky optimální, jestliže  $T$  jde do nekonečna.

### **Poděkování**

Výzkumný projekt je podporován grantem č. P402/12/G097 Grantové agentury České republiky a grantem č. IGA F4/57/2017, Fakulty informatiky a statistiky, VŠE, Praha.

### **Použitá literatura**

1. BITRAN, G., CALDENTEY, R. 2003. An Overview of Pricing Models for Revenue Management. *Manufacturing & Service Operations Management* 5, 203-229.
2. COOPER, W. L. 2002. Asymptotic Behavior of an Allocation Policy for Revenue Management Problem. *Operations Research* 50, 720-727.
3. FIALA, P. 2015. Revenue management – modely a analýzy. Professional Publishing, Praha.
4. KINCAID, W. M., DARLING, D. A. 1963. An Inventory Pricing Problem. *Journal of Mathematical Analysis Applications* 7, 183-208.

5. PHILLIPS, R. 2005. Pricing and Revenue Optimization. Stanford Business Books, Stanford, California.
6. TALLURI, K. T., VAN RYZIN, G. J. 2004. The Theory and Practice of Revenue Management. Kluwer Academic Publishers, Boston.

### **Kontaktní údaje**

Ing. Renata Majovská, PhD.

Vysoká škola finanční a správní, Fakulta ekonomických studií  
Estonská 500, 101 00 Praha 10  
Tel: (420 2) 10 088 831  
email: renata.majovska@gmail.com

Prof. RNDr. Ing. Petr Fiala, CSc., MBA

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky  
Nám. W. Churchilla 4, 130 67 Praha 3  
Tel: (420 2) 24 095 447  
email: pfiala@vse.cz

# PRODUCTION DATA IN THE SOCIAL ACCOUNTING MATRIX<sup>1</sup>

## PRODUKČNÉ ÚDAJE V MATICI SPOLOČENSKÉHO ÚČTOVNÍCTVA

*Veronika Mit'ková*

### Abstract

The paper deals with the social accounting matrix construction as a database for the computable general equilibrium models. The matrix for the Slovak Republic related to the remaining 27 European Union states and the aggregated rest of the world is presented. The matrix is based on the Global Trade Analysis Project data for the least available year, 2011. The second part of the paper is devoted to the production analysis, sectoral demandingness on the intermediates, factors of production and identifies the “workhorses” of the economy.

*Keywords:* social accounting matrix, intermediate inputs, production

### Abstrakt

Článok sa venuje konštrukcii matice spoločenského účtovníctva ako databáze modelov všeobecnej ekonomickej rovnováhy. Prezentovaná je matica pre Slovenskú republiku vo vzťahu k spoločenstvu ostávajúcich 27 krajín Európskej únie a agregovanému celku krajín zvyšku sveta. Matica je skonštruovaná na základe dát Global Trade Analysis Project za posledné dostupné obdobie, ktorým je rok 2011. V ďalšej časti sa článok zaoberá produkčnou analýzou a pojednáva o náročnosti jednotlivých odvetví z hľadiska medzivstupov a výrobných faktorov a identifikuje „ťahúňov“ ekonomiky.

*Kľúčové slová:* matica spoločenského účtovníctva, medzivstupy, produkcia

## 1 THE SOCIAL ACCOUNTING MATRIX STRUCTURE

The paper deals with the social accounting matrix (SAM) for Slovak republic, constructed on the basis of the Global Trade Analysis Project data (GTAP). The SAM represents monetary flows between activities, production sectors, institutions and the rest of the world. It usually serves as database for a computable general equilibrium model, input-output analysis or structural analysis. The basic features of the SAM are a) square matrix, b) column accounts record spending, c) row accounts record income, d) totals of the rows and corresponding columns are equal for each agent, e) each cell is an expenditure of an agent as well as a source of income to an agent. The structure of the SAM is inspired by Burfisher (2017), modified in order of columns with extracting some types of taxes (Table 1). The structure clearly shows the main macroeconomic aggregates such aggregate demand, aggregate supply, factors income, tax revenues, aggregate household and government income, foreign exchange inflow and outflow, savings and investments, aggregate supply, gross domestic production, factor expenditure, tax expenditure, aggregate expenditure, private and government consumption expenditure, gross investment expenditure. In the next chapter we discuss the identities based on the data structure.

---

<sup>1</sup> This paper research was supported by the Slovak Research and Development Agency, project No. APVV-14-0020: Pareto Optimal Interactions Modeling in the Financial and Business Adaptive Processes and by the the Slovak Ministry of Education, Science and Sport by grant VEGA 1/0340/16: “Modeling of financing and cooperation efficiency of the Slovak economy sectors and regions.”

**Table 1: Structure of the social accounting matrix (Burfisher, 2017, modified)**

	commodities		production activities	factors	taxes	final demand		regional household	final demand					total
	import variety	domestic variety				trade margins	ROW		private household	production tax	direct tax	government	investment	
imported commodity			demand for imported intermediates						demand for imports			demand for imports	demand for imports	aggregate demand
domestic commodity			demand for domestic intermediates			export of trade margins	exports		demand for domestic			demand for domestic	demand for domestic	
production activity		domestic production												domestic sales
factors			factor payments											factor income
taxes	import tariff	export tax	taxes on output, factor use, inputs	income tax					sales tax			sales tax	sales tax	tax revenue
trade margins	trade margins on imports													foreign exchange outflow
ROW	imports													
regional household				net factor income	tax revenues					production tax	direct tax			aggregate income
private household								household income						private hhd income
prod. tax			production tax											prod. tax
direct tax				direct tax										direct tax
government								gov't income						gov't income
savings				depreciation		foreign savings	foreign savings	domestic savings						savings
total	aggregate supply		gross domestic production	factor expenditure	tax expenditure	foreign exchange inflow		aggregate expenditure	private consumption expenditure	production tax	direct tax	gov't consumption expenditure	gross investment expenditure	

## 1.1 Aggregation of the Data Structure

The structure of the database is widely discussed in Narayanan, McDougall (2015). For the purpose of the current research, the database is aggregated to three regions, eleven sectors and five production factors.

The regional aggregation consists of three regions: Slovakia, 27 states of the European Union and 112 states of the rest of the world. The factors of production are aggregated to five groups: land, unskilled labor (includes clerks, service/shop workers, agricultural and unskilled), skilled labor (includes technicians, assoc. professional, officials and managers), capital and natural resources. The sectoral aggregation is divided into eleven groups as shows the Table 2.

**Table 2: Sectoral aggregation in SAM (Narayanan, McDougall, 2015)**

No.	aggregated	commodity/activity
T1	Auto	motor vehicles and parts, transport equipment nec
T2	Grains and Crops	paddy rice, wheat, cereal grains, vegetables, fruit, nuts, oil seeds, sugar cane, sugar beet, plant-based fibers, crops nec
T3	Livestock and Meat Products	cattle, sheep, goats, horses, animal products nec, raw milk, wool, silk-worm cocoons
T4	Mining and Extraction	forestry, fishing, coal, oil, colas, minerals nec
T5	Processed Food	sugar, food products nec, beverages and tobacco products
T6	Textiles and Clothing	textiles, wearing apparel
T7	Light Manufacturing	leather products, wood products, paper products, publishing, metal products, manufactures nec
T8	Heavy Manufacturing	petroleum, coal products, chemical, rubber, plastic products, mineral products nec, ferrous metals, metals nec, electronic equipment, machinery and equipment nec
T9	Utilities and Construction	electricity, gas manufacture, distribution, water, construction
T10	Transport and Communication	trade, transport nec, sea transport, air transport, communication
T11	Other Services	financial services nec, insurance, business services nec, recreation and other services, public administration, defense, health, education, dwellings

## 1.2 Database for Slovakia

The social accounting matrix for Slovak Republic was constructed according to the structure presented. The data come from the Global Trade Analysis Project (GTAP) for the latest available year, 2011. The data are in billions of US dollars. In the next chapters the paper deals with the production analysis and identifies the aggregate demand aggregates.

## 2 THE PRODUCTION ANALYSIS

The aggregated demand is the sum of the private households and the government demand, the investments, the demand for intermediate commodities and exports (include trade/transport export margins) as shows (2.1). These may be produced domestically or may be imported abroad (2.2 – 2.4).

$$\left[ \begin{array}{c} \text{aggregate} \\ \text{demand} \end{array} \right] = \left[ \begin{array}{c} \text{demand for} \\ \text{intermediates} \end{array} \right] + \left[ \begin{array}{c} \text{private households} \\ \text{demand} \end{array} \right] + \left[ \begin{array}{c} \text{government} \\ \text{demand} \end{array} \right] + [\text{investments}] + [\text{exports}] \quad (2.1)$$

For imported commodities:

$$80,855 = 60,507 + 13,694 + 264 + 6,386 + 0$$

For domestic commodities:

$$251,920 = 102,463 + 36,987 + 17,826 + 15,942 + 78,701$$

Total:

$$332,775 = 162,970 + 50,681 + 18,090 + 22,328 + 78,701$$

## 2.1 The Aggregate Demand

The Slovak production clearly sets the highest demand for imported intermediates in the heavy manufactures sector (38,199) and in the automotive sector (9,253), Table 3. For complexity, the automotive is a part of the heavy manufacture sector. The demand for domestic intermediates, the production activity of the country, is beside those two (70,010 and 23,502) also high in other services sector (57,853), transport and communication (33,519), utility and construction sectors (28,963). One can see the strong production orientation on manufactures and weak in agricultural sectors. This fact lies on historical background and specialization on Czechoslovak production in armaments industry.

**Table 3: Aggregate demand, own calculations**

Aggregate demand	imported	domestic
12 d_T1Auto	9 253	23 502
13 d_T2GrainsCr	1 292	1 648
14 d_T3MeatLstk	907	5 686
15 d_T4Extracti	8 899	2 073
16 d_T5ProcFood	3 215	9 121
17 d_T6TextWapp	2 744	2 665
18 d_T7LightMnf	7 801	16 880
19 d_T8HeavyMnf	38 199	70 010
20 d_T9Util_Con	1 755	28 963
21 d_T10TransCo	3 061	33 519
22 d_T11OthServ	3 729	57 853

Let's look inside the heavy manufactures and automotive sectors. The heavy industry sector uses intermediates from abroad (16,038) as well as from domestic heavy industry production (15,956). This industry is strongly inter-industry oriented using 62.5% of imported and 61.4% of domestic production in the own industry. Significant is the demand for extraction intermediates from abroad (6,709).

Slovakia belongs to the leading countries in the automotive production in the world with the most produced automobiles: 119 for 1,000 of inhabitants (Slovak Investment and Trade Development Agency, 2017). What does it mean for the inter-industry trade in the country? The intermediate demand is the highest for the imported automotive (6,347) and heavy industry (4,659) commodities, followed by significantly lower domestic heavy industry (2,565) and automotive (2,659) intermediate commodities, Table 4.

**Table 4: Production activities, own calculations**

		production activities										
		23 a_T1Au	24 a_T2Gr	25 a_T3Me	26 a_T4Ext	27 a_T5Prc	28 a_T6Te	29 a_T7Lig	30 a_T8He	31 a_T9Util	32 a_T10Tr	33 a_T11O
commodity import	1 m_T1Auto	6347	5,55	4,91	8,03	13,8	1,23	51,9	313	60,9	241	68,1
	2 m_T2GrainsCr	0,233	54,2	72,5	2,32	232	2,28	2,58	11,1	4,01	89,9	64,8
	3 m_T3MeatLstk	0,158	0,669	139	0,821	11	5,33	14,7	5,2	1,07	21,6	32
	4 m_T4Extracti	32,8	9,5	21	42	42,5	12,9	130	6709	960	355	248
	5 m_T5ProcFood	1,77	4,04	528	4,86	656	1,54	23	50,5	21,2	188	198
	6 m_T6TextWapp	8,87	0,866	2,49	3,36	11,6	570	142	98,1	29,4	108	120
	7 m_T7LightMnf	282	18,1	76,3	26,9	206	43,7	2136	1534	790	517	730
	8 m_T8HeavyMnf	4659	207	187	86,4	552	137	1452	16038	2220	1841	1933
	9 m_T9Util_Con	29,6	4,07	6,25	9,73	13,9	5,57	58,2	303	351	102	235
	10 m_T10TransCo	104	3	4,49	16,2	20,6	12,5	56,4	259	47,1	1057	283
	11 m_T11OthServ	71,3	19,7	29,2	11,2	88,8	17,1	117	353	230	523	1176
total import	11536,73	326,695	1071,14	211,821	1848,2	809,15	4183,78	25673,9	4714,68	5043,5	5087,9	
domestic commodities	12 d_T1Auto	2659	2,13	1,91	3,71	5,38	0,546	28,5	164	37,5	154	38,7
	13 d_T2GrainsCr	0,24	33,8	106	3,63	227	1,05	2,73	6,17	5,71	55,3	57,3
	14 d_T3MeatLstk	0,911	1,22	679	3,28	400	0,66	7,12	14,7	7,51	118	141
	15 d_T4Extracti	1,85	2,22	3,87	264	15,1	0,554	587	410	147	52,3	58,9
	16 d_T5ProcFood	3,15	4,23	626	8,26	1478	2,27	31	68,8	28,6	337	360
	17 d_T6TextWapp	8,46	0,441	0,565	0,759	2,71	188	46,9	32	9,92	29,6	37,1
	18 d_T7LightMnf	310	16,9	114	28,4	335	56,4	2488	1384	820	733	1136
	19 d_T8HeavyMnf	2565	183	114	109	314	87,5	1413	15956	2965	3496	1226
	20 d_T9Util_Con	266	42,3	76,5	102	185	51	589	2387	4757	1137	2514
	21 d_T10TransCo	1010	63,6	243	147	796	124	943	2280	1708	8760	3579
	22 d_T11OthServ	561	203	271	118	820	154	1079	3299	2355	4839	11339
total domestic	7385,611	552,841	2235,845	788,039	4578,19	665,98	7215,25	26001,67	12841,24	19711,2	20487	
total	18922,34	879,536	3306,985	999,86	6426,39	1475,13	11399,03	51675,57	17555,92	24754,7	25574,9	

## 2.2 The Intermediate Demand

The highest share on the imported intermediates has the textiles and clothing sector with more than 70% of inter-industry demand, followed by the grains and crops sector demand for the heavy manufactures (63.4%) and heavy manufactures inter-industry demand (62.5%). The automotive sector sets the highest demand for abroad intermediates from the automotive (55.0%) and again heavy industry sector (40.4%), Table 5.

**Table 5: Imported intermediate shares, own calculations**

shares for intermedie	23 a_T1Au	24 a_T2Gr	25 a_T3Me	26 a_T4Ext	27 a_T5Prc	28 a_T6Te	29 a_T7Lig	30 a_T8He	31 a_T9Util	32 a_T10Tr	33 a_T11O
1 m_T1Auto	55,0%	1,7%	0,5%	3,8%	0,7%	0,2%	1,2%	1,2%	1,3%	4,8%	1,3%
2 m_T2GrainsCr	0,0%	16,6%	6,8%	1,1%	12,6%	0,3%	0,1%	0,0%	0,1%	1,8%	1,3%
3 m_T3MeatLstk	0,0%	0,2%	13,0%	0,4%	0,6%	0,7%	0,4%	0,0%	0,0%	0,4%	0,6%
4 m_T4Extracti	0,3%	2,9%	2,0%	19,8%	2,3%	1,6%	3,1%	26,1%	20,4%	7,0%	4,9%
5 m_T5ProcFood	0,0%	1,2%	49,3%	2,3%	35,5%	0,2%	0,5%	0,2%	0,4%	3,7%	3,9%
6 m_T6TextWapp	0,1%	0,3%	0,2%	1,6%	0,6%	70,4%	3,4%	0,4%	0,6%	2,1%	2,4%
7 m_T7LightMnf	2,4%	5,5%	7,1%	12,7%	11,1%	5,4%	51,1%	6,0%	16,8%	10,3%	14,3%
8 m_T8HeavyMnf	40,4%	63,4%	17,5%	40,8%	29,9%	16,9%	34,7%	62,5%	47,1%	36,5%	38,0%
9 m_T9Util_Con	0,3%	1,2%	0,6%	4,6%	0,8%	0,7%	1,4%	1,2%	7,4%	2,0%	4,6%
10 m_T10TransCo	0,9%	0,9%	0,4%	7,6%	1,1%	1,5%	1,3%	1,0%	1,0%	21,0%	5,6%
11 m_T11OthServ	0,6%	6,0%	2,7%	5,3%	4,8%	2,1%	2,8%	1,4%	4,9%	10,4%	23,1%

The demand for the domestically produced intermediates is the highest in the heavy manufactures with 61.4% inter-industry share. The sector of the other services demands 55.3% share of the own sector. Transport and communication sets 44.4% demand in the inter-industry. Overall holds true that the intra-industry shares in all Slovak sectors are very high, with values above 28% (with expect of grains and crops).Automotive sets the highest demand in the own sector and heavy manufactures, about 35%, Table 6.

**Table 6: Domestic intermediate shares, own calculations**

	23 a_ T1Au	24 a_ T2Gr	25 a_ T3Me	26 a_ T4Ext	27 a_ T5Prc	28 a_ T6Te	29 a_ T7Lig	30 a_ T8He	31 a_ T9Util	32 a_ T10Tr	33 a_ T11O
12 d_ T1Auto	36,0%	0,4%	0,1%	0,5%	0,1%	0,1%	0,4%	0,6%	0,3%	0,8%	0,2%
13 d_ T2GrainsCr	0,0%	6,1%	4,7%	0,5%	5,0%	0,2%	0,0%	0,0%	0,0%	0,3%	0,3%
14 d_ T3MeatLstk	0,0%	0,2%	30,4%	0,4%	8,7%	0,1%	0,1%	0,1%	0,1%	0,6%	0,7%
15 d_ T4Extracti	0,0%	0,4%	0,2%	33,5%	0,3%	0,1%	8,1%	1,6%	1,1%	0,3%	0,3%
16 d_ T5ProcFood	0,0%	0,8%	28,0%	1,0%	32,3%	0,3%	0,4%	0,3%	0,2%	1,7%	1,8%
17 d_ T6TextWapp	0,1%	0,1%	0,0%	0,1%	0,1%	28,2%	0,7%	0,1%	0,1%	0,2%	0,2%
18 d_ T7LightMnf	4,2%	3,1%	5,1%	3,6%	7,3%	8,5%	34,5%	5,3%	6,4%	3,7%	5,5%
19 d_ T8HeavyMnf	34,7%	33,1%	5,1%	13,8%	6,9%	13,1%	19,6%	61,4%	23,1%	17,7%	6,0%
20 d_ T9Util_Con	3,6%	7,7%	3,4%	12,9%	4,0%	7,7%	8,2%	9,2%	37,0%	5,8%	12,3%
21 d_ T10TransCo	13,7%	11,5%	10,9%	18,7%	17,4%	18,6%	13,1%	8,8%	13,3%	44,4%	17,5%
22 d_ T11OthServ	7,6%	36,7%	12,1%	15,0%	17,9%	23,1%	15,0%	12,7%	18,3%	24,5%	55,3%

## 2.3 The Taxes

The total values of taxes paid by the sectors are described in Table 7. In the two agricultural sectors, grains and crops sector with the livestock and meat products, the total value of subsidies are higher than the sum of the taxes paid. The most important tax payer is the sector of other services and the heavy manufactures sector. The total value of taxes paid as the share on the production value is the highest in the mining and extraction, textiles and clothing and the other services sectors, about 9 – 10%. The most significant subsidies receiver related to the value of production is the grains and crops sector.

**Table 7: Taxes paid, own calculations**

Taxes paid		
23 a_ T1Auto	503	2,14%
24 a_ T2GrainsCr	135	-0,21%
25 a_ T3MeatLstk	53	-0,93%
26 a_ T4Extracti	194	0,33%
27 a_ T5ProcFood	457	0,01%
28 a_ T6TextWapp	251	0,40%
29 a_ T7LightMnf	876	0,19%
30 a_ T8HeavyMnf	2 680	0,83%
31 a_ T9Util_Con	1 403	0,84%
32 a_ T10TransCo	1 776	0,30%
33 a_ T11OthServ	5 234	0,05%

## 2.4 The Production Factors Income

The sector payments to the production factors, taxes excluded, are clearly the highest in the other services sector. The highest share on the factor income has the capital, again from the other services sector, Table 8.

**Table 8: Net factors income, own calculations**

Factors income	23 a_ T1Au	24 a_ T2Gr	25 a_ T3Me	26 a_ T4Ext	27 a_ T5Prc	28 a_ T6Te	29 a_ T7Lig	30 a_ T8He	31 a_ T9Util	32 a_ T10Tr	33 a_ T11O
34 Land	-	423	344	-	-	-	-	-	-	-	-
35 UnSkLab	560	154	294	191	518	284	983	2 598	1 607	919	2 057
36 SkLab	417	201	296	171	386	212	732	1 937	1 085	1 271	8 136
37 Capital	2 274	125	1 471	383	727	233	2 421	6 310	7 032	4 791	16 923
38 NatRes	-	-	-	160	-	-	-	-	-	-	-

## 3 CONCLUSIONS

The paper deals with the production analysis of the Slovak Republic, based on the 2011 data. The production sectors are aggregated to eleven groups with emphasis to the automotive



sector. The “workhorses” of the economy were identified in the sectors of heavy manufactures and other services. The analysis also proved the strong intra-industry orientation of the production, mainly in the domestic commodities intermediate use.

The future research plans to build on these basic findings and spread the analysis on the demand side of the economy.

### References

1. Automotive Sector in Slovakia, 2017. Slovak Investment and Trade Development Agency. Retrieved from <http://www.sario.sk/sites/default/files/data/sario-automotive-sector-in-slovakia-2017-08-17-web.pdf>
2. BURFISHER, M. 2016. Introduction to Computable General Equilibrium Models. Cambridge: Cambridge University Press.
3. NARAYANAN, B., MCDUGALL, R. 2015. Chapter 2: Guide to the GTAP Data Base (Center for Global Trade Analysis). Purdue University, West Lafayette, IN: Global Trade Analysis Project (GTAP). Retrieved from [https://www.gtap.agecon.purdue.edu/resources/res\\_display.asp?RecordID=4819](https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=4819)

### Kontaktné údaje

Ing. Veronika Miťková, PhD.

Ústav ekonómie, Fakulta sociálnych a ekonomických vied, Univerzita Komenského v Bratislave

Mlynské luhy 4, 821 05 Bratislava

email: [mitkova@fses.uniba.sk](mailto:mitkova@fses.uniba.sk)

# ALGEBRA VIACFAKTOROVEJ ÚLOHY VÝBERU PORTFÓLIA<sup>1</sup>

## MULTIFACTOR PORTFOLIO PROBLEM ALGEBRA

*Vladimír Mlynarovič*

### Abstrakt

Článok prezentuje rozvinutie modelu oceňovania kapitálových aktív do koncepcie viacfaktorového modelu pre úlohu výberu portfólia z rizikových aktív a úlohu doplnenú o bezrizikové aktívum. Výsledkom je algebrické riešenie zodpovedajúcich úloh a rozvinutie vety o vydelení dvoch fondov do jej viacfaktorovej verzie.

*Kľúčové slová:* dodatočný rizikový faktor, dotyčnicové portfólio, zaisťovacie portfólio

### Abstract

The paper presents extensions of the capital asset model into a concept of multifactor model for portfolio selection model with risky assets and model extended with a riskless asset. As a result we have algebraic solutions of corresponding problems and extensions of the theorem of two funds separation into its multifactor version.

*Keywords:* extra risk factor, tangency portfolio, hedging portfolio

## 1 ÚVOD

Model oceňovania kapitálových aktív (CAPM) bol v minulosti považovaný za nástroj, ktorý poskytuje dobré vysvetlenie toho, prečo sú priemerné výnosy niektorých akcií, portfólií, fondov alebo stratégií vyššie ako u iných. Teraz sa ukazuje, že priemerné výnosy mnohých investičných príležitostí nemožno vysvetliť prostredníctvom CAPM a jeho miesto zaujímajú „viacfaktorové modely“. Uvažovalo sa, že priemerné výnosy vzájomných fondov sú dobré opísané prostredníctvom CAPM. Teraz možno pozorovať, že fondy môžu sledovaním palety investičných „štýlov“ zarábať priemerné výnosy, ktoré nie sú vysvetlené CAPM, t.j. nie sú spojené s trhovým rizikom.

Rozvinutia CAPM do formy viacfaktorových modelov, v ktorých sa vyššie priemerné výnosy aktív spájajú, ako doplnok k pohybu trhu ako celku, s tendenciou pohybu aj s inými rizikovými faktormi, dominujú pri opisovaní atribútov výkonnosti a pri vysvetľovaní priemerných výnosov týchto aktív. Revolúcia v nazeraní finančných ekonómov na svet teda vyústila do záveru, že existujú stratégie, ktorých výsledkom sú vysoké priemerné výnosy bez vysokých faktorov beta, t.j. bez silnej tendencie k tomu, aby sa výnosové stratégie pohybovali hore a dole s pohybom trhu ako takého. Viacfaktorové modely vytlačili model oceňovania kapitálových aktív (CAPM) pri opisovaní týchto fenoménov.

---

<sup>1</sup> This paper research was supported by the Slovak Research and Development Agency, project No. APVV-14-0020: “Pareto Optimal Interactions Modeling in the Financial and Business Adaptive Processes”, and by the Slovak Ministry of Education, Science and Sport, project No. VEGA 1/0340/16: “Modelovanie efektívnosti financovania a kooperácie sektorov a regiónov slovenskej ekonomiky.”

Teoretické základy nového prístupu ku konštrukcii portfóliových odporúčaní, ktorý rozvíja vyššie charakterizovaný prístup klasickej teórie portfólia v tom zmysle, že odvíja konštrukciu odporúčaní nie len od trhového rizika, ale aj od dodatočných rizikových faktorov boli vybudované a prezentované v klasickej publikácii *R. C. Mertona* (1973). Klasický výsledok o vydelení dvoch fondov bol rozvinutý do vety o vydelení  $m + 2$  fondov, kde  $m$  je počet dodatočných rizikových faktorov, ktoré sú predmetom záujmu investora. Nasledujúca analytická konštrukcia, ktorá opisuje matematické aspekty úlohy výberu portfólia s dodatočným rizikovým faktorom sumarizuje *Famov* (1996) a *Cochraneho* (1999) prístup.

## 2 PORTFÓLIO RIZIKOVÝCH AKTÍV

Úlohou je minimalizovať rozptyl portfólia pri danej hodnote očakávaného výnosu portfólia a jeho beta na nejaký dodatočný rizikový faktor. Nech

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}, \mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}, \mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_{1,F} \\ \beta_{2,F} \\ \vdots \\ \beta_{n,F} \end{bmatrix}$$

potom pre výnos portfólia platí

$$r_p = \mathbf{w}^T \mathbf{r}$$

pre váhy musí byť splnená podmienka

$$1 = \mathbf{e}^T \mathbf{w}$$

a očakávaný výnos  $E_p$  portfólia je rovný

$$E_p = E(r_p) = E(\mathbf{w}^T \mathbf{r}) = \mathbf{w}^T E(\mathbf{r}) = \mathbf{w}^T \mathbf{E}$$

kde posledná z rovníc len zjednodušuje označenie. Nech pre beta portfólia na dodatočný rizikový faktor platí

$$\beta_p = \mathbf{w}^T \boldsymbol{\beta}$$

Rozptyl výnosu portfólia je

$$\text{var}(r_p) = \mathbf{w}^T \mathbf{C} \mathbf{w}$$

kde  $\mathbf{C}$  je variačno – kovariačna matica výnosov. Úloha má teraz tvar

$$\min \frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w}$$

za podmienok

(2.1)

$$\mathbf{w}^T \mathbf{E} = E_p, \quad \mathbf{w}^T \mathbf{e} = 1, \quad \mathbf{w}^T \boldsymbol{\beta} = \beta_p$$

Lagrangeova funkcia má tvar

$$\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} - \lambda_0 (\mathbf{w}^T \mathbf{E} - E_P) - \lambda_1 (\mathbf{w}^T \mathbf{e} - 1) - \lambda_2 (\mathbf{w}^T \boldsymbol{\beta} - \beta_P)$$

a podmienky prvého rádu vzhľadom k  $\mathbf{w}$  dávajú

$$\mathbf{w} = \mathbf{C}^{-1} (\mathbf{E} \lambda_0 + \mathbf{e} \lambda_1 + \boldsymbol{\beta} \lambda_2) = \mathbf{C}^{-1} \mathbf{B} \boldsymbol{\lambda}$$

kde

$$\mathbf{B} = [\mathbf{E} \ \mathbf{e} \ \boldsymbol{\beta}]$$

$$\boldsymbol{\lambda} = [\lambda_0, \lambda_1, \lambda_2]^T$$

$$\boldsymbol{\delta} = [\mu, 1, \beta_P]^T$$

Po dosadení tejto hodnoty  $\mathbf{w}$  do ohraničujúcich rovníc vyplývajúcich zo zostávajúcich podmienok prvého rádu v tvare

$$\mathbf{B}^T \mathbf{w} = \boldsymbol{\delta}$$

postupne dostaneme

$$\mathbf{B}^T \mathbf{C}^{-1} \mathbf{B} \boldsymbol{\lambda} = \boldsymbol{\delta}$$

$$\boldsymbol{\lambda} = (\mathbf{B}^T \mathbf{C}^{-1} \mathbf{B})^{-1} \boldsymbol{\delta}$$

odkiaľ pre optimálny vektor váh rizikových aktív platí

$$\mathbf{w} = \mathbf{C}^{-1} \mathbf{B} (\mathbf{B}^T \mathbf{C}^{-1} \mathbf{B})^{-1} \boldsymbol{\delta} \quad (2.2)$$

a optimálny rozptyl portfólia je potom rovný

$$\text{var}(r_P) = \mathbf{w}^T \mathbf{C} \mathbf{w} = \boldsymbol{\delta}^T (\mathbf{B}^T \mathbf{C}^{-1} \mathbf{B})^{-1} \boldsymbol{\delta} \quad (2.3)$$

alebo, po rozpísaní maticovej formulácie,

$$\text{var}(r_P) = [\mu \ 1 \ \beta_P] (\mathbf{B}^T \mathbf{C}^{-1} \mathbf{B})^{-1} \begin{bmatrix} \mu \\ 1 \\ \beta_P \end{bmatrix}$$

### 3 BEZRIZIKOVÉ AKTÍVUM

Teraz predpokladajme, že okrem  $n$  rizikových aktív existuje bezrizikové aktívum so sadzbou  $r_f$ . Potom  $\gamma_i = E_i - r_f$  je dodatočný očakávaný výnos rizikového aktíva  $i$ ,  $i = 1, \dots, n$  a  $\boldsymbol{\gamma}$  nech je  $n$  – rozmerný vektor dodatočných očakávaných výnosov rizikových aktív. Ak  $\gamma_P$  je

požadovaný dodatočný očakávaný výnos portfólia špecifikovaný investorom, potom sa úloha výberu portfólia (5.15) transformuje na úlohu

$$\min \frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w}$$

za podmienok

$$(3.1)$$

$$\mathbf{w}^T \mathbf{r} = \gamma_P, \quad \mathbf{w}^T \boldsymbol{\beta} = \beta_P$$

pričom čitateľ si už určite všimol, že koeficient 1/2 v účelovej funkcii slúži len na zjednodušenie riešenia úlohy prostredníctvom nutných a postačujúcich podmienok prvého rádu, podľa ktorých pre *Lagrangeovu funkciu* v tvare

$$L(\mathbf{w}, \lambda_0, \lambda_1) = \frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} - \lambda_0 (\mathbf{w}^T \boldsymbol{\gamma} - r_P) - \lambda_1 (\mathbf{w}^T \boldsymbol{\beta} - \beta_P)$$

platí

$$\mathbf{C} \mathbf{w} - \lambda_0 \mathbf{r} - \lambda_1 \boldsymbol{\beta} = 0$$

$$\mathbf{w}^T \boldsymbol{\gamma} - r_P = 0$$

$$\mathbf{w}^T \boldsymbol{\beta} - \beta_P = 0$$

odkiaľ

$$\mathbf{w} = \mathbf{C}^{-1}(\lambda_0 \boldsymbol{\gamma} + \lambda_1 \boldsymbol{\beta}) = \mathbf{C}^{-1} \begin{bmatrix} \boldsymbol{\gamma} & \boldsymbol{\beta} \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \end{bmatrix} = \mathbf{C}^{-1} \mathbf{D} \boldsymbol{\lambda}$$

$$\mathbf{D}^T \mathbf{w} = \begin{bmatrix} \gamma_P \\ \beta_P \end{bmatrix}$$

a ďalej

$$\mathbf{D}^T \mathbf{C}^{-1} \mathbf{D} \boldsymbol{\lambda} = \mathbf{G} \boldsymbol{\lambda} = \begin{bmatrix} \gamma_P \\ \beta_P \end{bmatrix}$$

$$\boldsymbol{\lambda} = \mathbf{G}^{-1} \begin{bmatrix} \gamma_P \\ \beta_P \end{bmatrix}$$

na základe čoho môžeme optimálny vektor váh rizikových aktív vyjadriť v tvare

$$\mathbf{w} = \mathbf{C}^{-1} \begin{bmatrix} \boldsymbol{\gamma} & \boldsymbol{\beta} \end{bmatrix} \mathbf{G}^{-1} \begin{bmatrix} \gamma_P \\ \beta_P \end{bmatrix} \quad (3.2)$$

a optimálny rozptyl výnosov v tvare

$$\sigma_P^2 = \begin{bmatrix} \gamma_P & \beta_P \end{bmatrix} \mathbf{G}^{-1} \begin{bmatrix} \boldsymbol{\gamma} & \boldsymbol{\beta} \end{bmatrix}^T \mathbf{C}^{-1} \mathbf{C} \mathbf{C}^{-1} \begin{bmatrix} \boldsymbol{\gamma} & \boldsymbol{\beta} \end{bmatrix} \mathbf{G}^{-1} \begin{bmatrix} \gamma_P \\ \beta_P \end{bmatrix} = \begin{bmatrix} \gamma_P & \beta_P \end{bmatrix} \mathbf{G}^{-1} \begin{bmatrix} \gamma_P \\ \beta_P \end{bmatrix} \quad (3.3)$$

Teraz, aby sme ukázali ako možno v tejto konkrétnej situácii, v ktorej sa uvažuje práve s jedným dodatočným rizikovým faktorom, vymedziť 3 fondy ( $m + 2 = 1 + 2$ ), prostredníctvom ktorých možno opísať hranicu množiny investičných príležitosti, preskúame detailne vzťah (3.2). Označme prvky matice  $\mathbf{G}$  nasledovne

$$\mathbf{G} = \begin{bmatrix} \gamma^T \mathbf{C}^{-1} \gamma & \gamma^T \mathbf{C}^{-1} \beta \\ \beta^T \mathbf{C}^{-1} \gamma & \beta^T \mathbf{C}^{-1} \beta \end{bmatrix} = \begin{bmatrix} v & s \\ s & u \end{bmatrix}$$

Ak maticu  $\mathbf{G}^{-1}$  teraz vyjadríme v tvare

$$\mathbf{G}^{-1} = \frac{1}{vu - s^2} \begin{bmatrix} u & -s \\ -s & v \end{bmatrix}$$

potom pre vektor  $\mathbf{w}$  podľa vzťahu (3.2) platí.

$$\mathbf{w} = \frac{1}{vu - s^2} [(u\gamma_P - s\beta_P)\mathbf{C}^{-1}\gamma + (v\beta_P - s\gamma_P)\mathbf{C}^{-1}\beta] \quad (3.4)$$

Na základe analýzy úlohy výberu portfólia pri predpoklade o existencii bezrizikového aktíva (Mlynarovič, 2001), pre váhy  $\mathbf{w}_T$  dotyčnicového portfólia platí

$$\mathbf{w}_T = \frac{1}{\mathbf{e}^T \mathbf{C}^{-1} \gamma} \mathbf{C}^{-1} \gamma$$

Ak definujeme vektor váh  $\mathbf{w}_H$  zaist'ovacieho portfólia, ktorým sa investor zaist'uje vzhľadom k uvažovanému dodatočnému rizikovému faktoru, v tvare

$$\mathbf{w}_H = \frac{1}{\mathbf{e}^T \mathbf{C}^{-1} \beta} \mathbf{C}^{-1} \beta$$

potom vzťah (3.4) možno vyjadriť v tvare

$$\mathbf{w} = \alpha_T \mathbf{w}_T + \alpha_H \mathbf{w}_H \quad (3.5)$$

kde pre váhy  $\alpha_T$  a  $\alpha_H$  platí

$$\alpha_T = \frac{u\gamma_P - s\beta_P}{vu - s^2} \mathbf{e}^T \mathbf{C}^{-1} \gamma, \quad \alpha_H = \frac{v\beta_P - s\gamma_P}{vu - s^2} \mathbf{e}^T \mathbf{C}^{-1} \beta$$

Keďže tieto váhy závisia od hodnôt  $\gamma_P$  a  $\beta_P$  špecifikovaných individuálnym investorom, vyjadrujú preferencie individuálneho investora v priestore očakávaný výnos – rozptyl – dodatočný rizikový faktor. Ak označíme ako  $w_0$  váhu bezrizikového aktív v portfóliu a pripomenieme, že

$$w_0 = 1 - \sum_{i=1}^n w_i$$

potom trojica  $w_0$ ,  $\mathbf{w}_T$  a  $\mathbf{w}_H$  hľadané opisuje tri fondy.

**Použitá literatúra**

1. COCHRANE, J. H. 1999. Portfolio advice for multifactor world. National bureau of economic research.
2. FAMA, E. F. 1996. Multifactor portfolio efficiency and multifactor asset pricing,. *Journal of Financial and Quantitative Analysis*, Vol. 31, No. 4, December, pp. 441.465.
3. MERTON, R. C. 1973. An intertemporal capital asset pricing model,. *Econometrica*, Vol. 41, No. 5, pp. 867-887.
4. MLYNAROVIC, V. 2001. Finančné investovanie. Bratislava, IURA Edition.

**Kontaktné údaje**

doc. Ing. Vladimír Mlynarovič, CSc.

Univerzita Komenského v Bratislave, Fakulta sociálnych a ekonomických vied

Mlynské luhy 4 1/b, 821 05 Bratislava

Tel: +421 2 20 66 98 60

email: vladimír.mlynarovic@fses.uniba.sk

# TECHNOLÓGIA SPRACOVANIA AUTOMOBILOVÉHO ODPADU A ŽIVOTNÝ CYKLUS AUTOMOBILOV<sup>1</sup>

## AUTOMOBILE WASTE PROCESSING TECHNOLOGY AND LIFE CYCLES OF AUTOMOBILES

*Juraj Pekár*

### **Abstrakt**

V príspevku si rozoberieme technológiu, ktorá sa využíva pri recyklovaní automobilov a životný cyklus automobilov, od výroby až po možné zhodnotenie materiálov a opätovné využitie jednotlivých komponentov, konkrétne postup recyklácie a ďalšie spracovanie materiálov získaných z jednotlivých častí vyradených automobilov. Trend nárastu využívania automobilov vytvára vážne obavy z environmentálnych následkov, pričom jedným zo spôsobov zníženia uvedených dopadov na životné prostredie je nastavenie procesov recyklácie starých automobilov.

*Kľúčové slová: recyklácia, automobil*

### **Abstract**

In the paper, we will discuss the technology used for car recycling and the life cycle of cars, from production to the possible recovery of materials and reuse of individual components, namely the recycling process and the further processing of materials obtained from individual parts of discarded cars. The growing trend of car use increases concerns of its environmental impact. One way to reduce these environmental impacts is to set up the recycling process for old cars.

*Keywords: recycling, car*

## **1 ÚVOD**

Rastúca prosperita viedla k obrovskému používaniu automobilov vo svete. Každý rok sa v Európskej únii vyprodukuje 8 až 9 miliónov ton odpadu pozostávajúceho z vyradených vozidiel. Automobilový priemysel Slovenskej republiky, zásluhou výrobných závodov Volkswagen, PSA Peugeot-Citroen a KIA Motors, sa dostal na prvú priečku slovenského priemyslu. Avšak nemali by sme zanedbať negatíva spojené s rozvojom automobilovej výroby a následnej spotreby automobilov. V rámci environmentálnej stratégie Slovenskej republiky, pri neustále sa zhoršujúcom životnom prostredí a súčasnom zvýšení produkcie nových automobilov, je dôležité venovať vysokú pozornosť pri výbere vhodných materiálov použitých pri výrobe ako aj neškodnú recykláciu alebo demontáž nepojazdného automobilu. Smernica Európskeho parlamentu stanovila postup procesu likvidácie vozidla, aby bolo dosiahnuté zhodnotenie na vysokej ekologickej úrovni. Stanovila, že 95 % hmotnosti vozidla musí byť opätovne využitých alebo zhodnotených. Direktíva určuje jasne ciele opätovného

---

<sup>1</sup> Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA No. 1/0351/17 Application of selected models of game theory to solve some economic problems of Slovakia

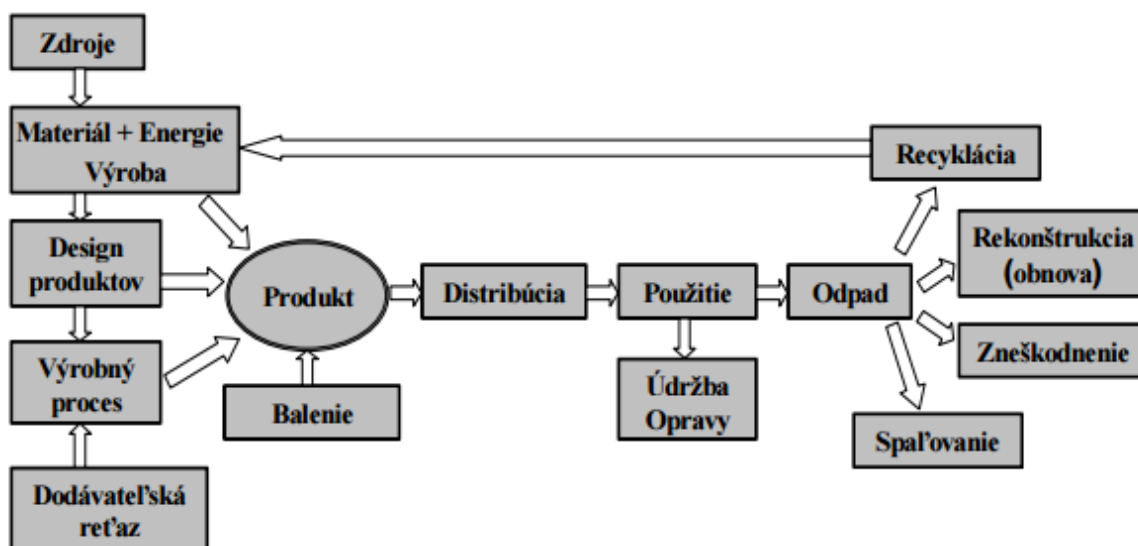


využitia, recyklácie a obnovy áut, a ich komponentov a vyvíja tlak na výrobcov, aby už pri výrobe nových vozidiel zohľadnili recyklovateľnosť.

## 2 ŽIVOTNÝ CYKLUS AUTOMOBILU

Každá časť životného cyklu automobilu môže negatívne vplyvať na životné prostredie. Pri výrobe použitá technológia ako aj materiály majú výrazný vplyv na fázu zneškodňovania a rozoberania. Využívanie automobilu je ďalšou fázou jeho životného cyklu. Riziko vzhľadom na životné prostredie tu predstavuje havária, rôzny únik prevádzkových kvapalín a emisie, ktoré vznikajú z premeny energie spaľovaním. Poslednou fázou je zneškodňovanie automobilu po dobe jeho životnosti. Je dôležité dodržiavať deštruktívnu a nedeštruktívnu demontáž jednotlivých súčiastok a častí automobilu a ich následné triedenie a recyklácia. Existujú pravidlá, ktoré napomáhajú pri navrhovaní nových automobilov. Vychádzajú z metodiky integrovanej výrobkovej politiky tzv. Integrated Product Policy, v skratke IPP. Ako prvé pravidlo je čistejšia výroba, v rámci ktorej sa preverujú technologické postupy, množstvá vyprodukovaných odpadov a spotreba jednotlivých energií. Nasleduje snaha znížiť spotrebu materiálov a súčasne znížiť aj počet rôznych druhov materiálov v jednom výrobku, čím nastáva redukcia hmotnosti. Eliminovanie negatívneho vplyvu na životné prostredie pri výrobe automobilov možno realizovať použitím recyklovateľných materiálov, nahrádzaním toxických a nebezpečných materiálov ekologickejšími variantami. Zníženie negatívnych vplyvov na životné prostredie možno dosiahnuť vyššou životnosťou, čiže jednoduchšou opraviteľnosťou a preventívnymi kontrolami za účelom udržiavať automobil v dobrom stave. Posledným pravidlom je zjednodušiť konečnú recykláciu automobilu tak, aby nebola zbytočne komplikovaná deštrukcia a jednoduchšie triedenie materiálov zabezpečiť jednotným systémom označovania materiálov.

Obrázok 1: Životný cyklus automobilu



Zdroj: <https://www.sjf.tuke.sk/transferinovacii/pages/archiv/transfer/13-2009/pdf/013-016.pdf>

Životný cyklus automobilu možno znázorniť prostredníctvom schémy-na obrázku 1.<sup>2</sup> Keďže nás zaujíma tvorba odpadov a ich následné spracovanie, zameriame sa na proces recyklácie.

<sup>2</sup> LEŠKOVÁ, A., a kol., 2009: *Finálna fáza životného cyklu automobilu*. [online] In Transfer inovácií 13/2009. Príspevok bol spracovaný v rámci riešenia projektu VEGA 1/0248/09 *Inovačné techniky pre navrhovanie a*

Recykláciu alebo inak povedané aj opätovné využitie môžeme vnímať aj ako spätnú logistiku. Spätnou logistikou vytvárame pohyb opotrebovaných alebo inak znehodnotených produktov na konci ich samotného životného cyklu. Alebo si môžeme pojem recyklácia zadefinovať ako postup vrátenia odpadu do výrobného cyklu na výrobu požadovaného výrobku s cieľom šetrenia surovinových zdrojov.<sup>3</sup> Recyklácia sa vytvára pomocou recyklačných technológií. Pod recyklačnými technológiami si môžeme predstaviť súbor na seba nadväzujúcich výrobných procesov.

V rámci životného cyklu automobilu sa zameriame na segmenty, kde je tvorený odpad určený na recyklovanie v najväčšej miere, či už v rámci použitia automobilu pri činnosti - údržba a opravy, alebo po skončení použitia automobilu. Z hľadiska technologického sa skúmajú možnosti ako optimálne recyklovať odpad na opätovne použiteľnú surovinu. V práci (Pekár a kol., 2017) riešia autori problematiku umiestnenia autorizovaných spracovateľov starých automobilov. Existuje však mnoho častí procesu recyklácie, kde by bolo vhodné použiť optimalizáciu k dosiahnutiu požadovaných cieľov smerujúcich k zníženiu environmentálnej záťaže.

### 3 TECHNOLOGIA SPRACOVANIA STARÝCH VOZIDIEL

Pri bežnej prevádzke je prevažne riešený problém recyklácie prevádzkových kvapalín a demontáže opotrebovaných súčiastok, ktoré sú následne zaslané na recykláciu. Uvedený proces realizuje každý servis individuálne, pričom sú stanovené pravidlá, akým spôsobom nakladať s odpadom. Komplikovanejšia je situácia pri procese likvidácie starých vozidiel, preto si v tejto časti opíšeme technológiu spracovania starých vozidiel.

Pri spracovaní vyradených automobilov sa používajú tieto druhy technológií: bezdemontážna technológia a predbežná demontáž. Návrh technológií, ktoré sa využívajú pri recyklácii, musia vychádzať z ekonomickej rentability, minimálneho vplyvu na životné prostredie, možnosti opätovného použitia recyklovaného materiálu. Pri *bezdemontážnej technológii* sa automobily celé zlisujú a následne sú prostredníctvom dopravníka dopravené do tzv. rotačného kladivového mlyna, kde sa tento zlisovaný materiál trhá na kusový šrot. Následne je tento kusový šrot vytriedený na kovové a nekovové materiály. Niektorý odpadový materiál je ťažko spracovateľný a tým pádom sa vyváža na skládky odpadu alebo do spaľovní.

Pri technológii, ktorá využíva *predbežnú demontáž* pri spracovaní starého automobilu, je prvotným krokom predbežná demontáž a následne sa prejde k jednotlivému drveniu v drvičoch. Na spracovanie sa jednotlivé vozidlá privážajú už zbavené súčiastok ako je motor, batérie, plášte pneumatík a iné diely. Starý automobil musí byť zbavený prevádzkových kvapalín ako je olej, benzín, brzdová kvapalina atď. Nižšie na obrázku 2 je zobrazená technológia spracovania.<sup>4</sup>

Pri spracovaní automobilu sa nezohľadňuje, či ide o auto v plnej výbave, ktorú dodal výrobca automobilu alebo, ktorú si držiteľ upravil sám v súlade s predpismi na údržbu alebo opravu motorových vozidiel. Firma, ktorá sa zaoberá recykláciou, by mala z hľadiska rentability zvoliť vhodný postup na spracovanie starých vozidiel. Vzhľadom na to, že nastanú situácie,

---

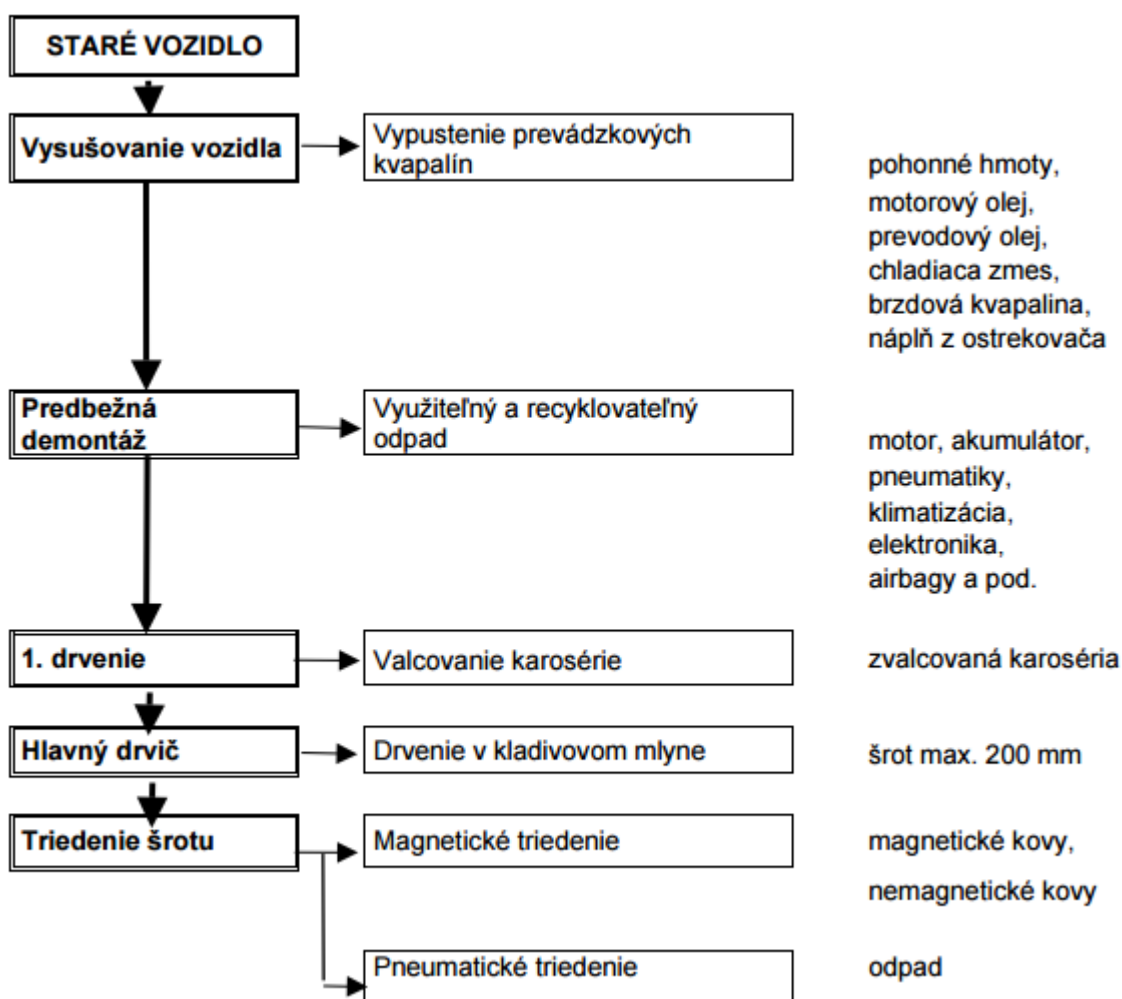
testovanie automobilových komponentov. Dostupné na internete: <<https://www.sjf.tuke.sk/transferinovacii/pages/archiv/transfer/13-2009/pdf/013-016.pdf>>.

<sup>3</sup> JUNÁK, J., 2012: *Recyklačné technológie 1. časť*. [Online]. In Košice 2012. Vydavateľstvo: Technická univerzita v Košiciach, Stavebná fakulta. Dostupné online: <<http://www.svf.tuke.sk/wp-content/uploads/2014/09/Ucebne-texty-Recyklačne-technologie-1.pdf>>. ISBN: 987-80-553-1254-5.

<sup>4</sup> KOVÁČ, B., a kol., 2009: *Ekologické spracovanie osobných automobilov po dobe životnosti*. [Online] In Doprava a spoje –elektronický časopis Fakulty prevádzky a ekonomiky dopravy a spojov Žilinskej univerzity v Žiline. Dostupné na internete: <<https://fpedas.uniza.sk/dopravaaspoje/2009/1/kovac.pdf>>. ISSN 1336-7676.

kedy zostatková hodnota vozidla je nižšia ako náklady spojené s jeho spracovaním. To platí väčšinou pre automobily staršie ako 15 rokov. Ekonomicky výhodnejšie na spracovanie sú novšie automobily, ktoré boli vyradené z evidencie na spracovanie z dôvodu napríklad dopravnej nehody. Tieto automobily predstavujú vysokú využiteľnosť z hľadiska náhradných dielov, aj keď hodnota týchto dielov je nižšia ako cena originálnych nových dielov. Vhodnou voľbou postupu spracovania automobilov firma dokáže znížiť časovú náročnosť, mzdové náklady alebo úroveň technického vybavenia.<sup>5</sup> Vhodnou separáciou a následným spracovaním môžu byť jednotlivé materiály použité ako vstupná surovina pre ďalšiu výrobu.

Obrázok 2 : Schéma technológie spracovania starých vozidiel



Zdroj: <https://fpedas.uniza.sk/dopravaaspoje/2009/1/kovac.pdf>

#### 4 LOKÁCIA SPRACOVATEĽOV ODPADOV Z AUTOMOBILOV

V predchádzajúcich častiach bol opísaný životný cyklus automobilu a technológia spracovania starých automobilov. Na základe obrázkov 1 a 2 možno získať informácie

<sup>5</sup>HLADNIŠ, J. a kol., 2008: *Logistický proces demontáže cestných vozidiel a ich ekonomická využiteľnosť*. Pardubice: Univerzita Pardubice 2008, Dopravná fakulta Jana Pernera. Elektronický odborný časopis o technológii, technice a logistice v doprave. Ročník 3. Číslo 8. 2008 [Online] Dostupné na internete: < [http://pernerscontacts.upce.cz/10\\_2008/Hladnis\\_Masek.pdf](http://pernerscontacts.upce.cz/10_2008/Hladnis_Masek.pdf) >. ISSN 1801 -674X.

o fázach, v ktorých vznikajú odpady z automobilov. Vo fáze používania je to pri údržbe a oprave automobilu a po životnosti pri jeho likvidácii. Pre dosiahnutie zvýšenia ekologickej likvidácie je nutné zabezpečenie ekonomicky a ekologicky primeranej siete spracovateľov odpadov. Vo fáze údržby a opravy uvedený proces realizujú servisné centrá, ktoré ekologicky likvidujú prevádzkové kvapaliny, ako aj nefunkčné časti vozidiel. Lokáciu uvedených centier nie je nutné optimalizovať, keďže zabezpečujú bežnú prevádzku vozidiel. Iná situácia je so zabezpečením recyklačných centier určených na likvidáciu starých vozidiel, keďže je technologicky náročnejšia (obrázok 2). V súčasnosti je evidovaných 46 spracovateľských zariadení určených na likvidáciu vyradených a nepojazdných dopravných prostriedkov. Na stanovenie optimálnej siete uvedených spracovateľov možno použiť optimalizačné prístupy zabezpečujúce ich dostupnosť pre všetkých potencionálnych zákazníkov. Ako vhodný nástroj možno použiť lokačné modely. Cieľ lokačných modelov vo všeobecnosti je optimálne umiestniť alebo rozmiestniť centrá, zariadenia, konkrétne všetko čo súvisí s územným plánovaním. Jednotlivé modely si vyžadujú pochopiť fungovanie skutočných situácií. Optimálne umiestnenie objektov závisí od vhodne zvolených optimalizačných metód, ktoré sú určené na riešenie uvedeného problému. Modely umiestnenia jednotlivých zariadení vychádzajú predovšetkým z koncepcie pokrytia obyvateľov obslužnými strediskami. Vo všeobecnosti môžeme uvažovať s dvomi typmi základných lokačných modelov. Prvý model je založený na pokrytí všetkých požiadaviek minimálnym počtom zariadení pri danej maximálnej vzdialenosti, druhý model je založený na maximalizácii pokrytia požiadaviek pri obmedzenom počte obslužných zariadení.<sup>6</sup>

## 5 ZÁVER

V rámci enviromentálnej stratégie Slovenskej republiky, pri neustále sa zhoršujúcom životnom prostredí a súčasnom zvýšení produkcie nových automobilov, je dôležité venovať vysokú pozornosť procesom recyklácie automobilov a odpadom vznikajúcim pri jeho použití. Príspevok je venovaný životnému cyklu automobilov, pričom sa sústreďuje na segmenty, v ktorých vznikajú odpady v procese využívania automobilu a taktiež po jeho znefunkčnení. V ďalšej časti bola rozobraná technológia spracovania starého automobilu, pričom sú uvedené konkrétne postupy recyklácie. V poslednej časti sú uvedené nástroje použiteľné na rozhodovanie o umiestnení spracovateľov odpadov z automobilov. Keďže dlhodobo pozorujeme nárast využívania automobilov, je potrebné eliminovať množstvo nerecyklovateľných odpadov, čo znižuje negatívne dopady na životné prostredie. Uvedený jav možno dosiahnuť analyzovaním a nastavovaním procesov výroby a recyklácie automobilov, pričom jedným zo spôsobov je využívanie optimalizačných modelov a metód. Príkladom pri procese recyklácie je optimalizácia počtu zberných miest (minimalizácia nákladov) pomocou lokačných modelov, ktoré zabezpečia lepšiu dostupnosť pre všetkých obyvateľov, čím ich motivujú ku ekologickej likvidácii starých automobilov.

### Použitá literatúra

1. HLADNIŠ, J. a kol., 2008: Logistický proces demontáže cestných vozidiel a ich ekonomická využiteľnosť. Pardubice: Univerzita Pardubice 2008, Dopravní fakulta Jana Pernera. Elektronický odborný časopis o technológii, technice a logistice v

---

<sup>6</sup>BREZINA, I., DUPAL, A., PEKÁR, J., 2011: *Zelená a reverzná logistika ako nástroj zefektívnenia spaľovania odpadu v Slovenskej republike*. [Online] In Ekonomický časopis, 59, 2011, č.2, s. 132-147. Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/0360/10 Model rozmiestnenia recyklačných miest pre Slovenskú republiku. Dostupné na internete: <<https://www.sav.sk/journals/uploads/0913103802%2011%20Brezina%20a%20kol.RS.pdf>>.

- doprave. Ročník 3. Číslo 8. 2008 [Online] Dostupné na internete: <[http://pnerscontacts.upce.cz/10\\_2008/Hladnis\\_Masek.pdf](http://pnerscontacts.upce.cz/10_2008/Hladnis_Masek.pdf)>. ISSN 1801 -674X.
2. JUNÁK, J., 2012: Recyklačné technológie 1. časť. [Online]. In Košice 2012. Vydavateľstvo: Technická univerzita v Košiciach, Stavebná fakulta. Dostupné online: <<http://www.svf.tuke.sk/wp-content/uploads/2014/09/Ucebne-texty-Recyklačne-technologie-1.pdf>>. ISBN: 987-80-553-1254-5.
  3. KOVÁČ, B., a kol., 2009: Ekologické spracovanie osobných automobilov po dobe životnosti. [Online] In Doprava a spoje –elektronický časopis Fakulty prevádzky a ekonomiky dopravy a spojov Žilinskej univerzity v Žiline. Dostupné na internete: <<https://fpedas.uniza.sk/dopravaaspoje/2009/1/kovac.pdf>>. ISSN 1336-7676.
  4. LEŠKOVÁ, A., a kol., 2009: Finálna fáza životného cyklu automobilu. [online] In Transfer inovácií 13/2009. Príspevok bol spracovaný v rámci riešenia projektu VEGA 1/0248/09 Inovačné techniky pre navrhovanie a testovanie automobilových komponentov. Dostupné na internete: <<https://www.sjf.tuke.sk/transferinovacii/pages/archiv/transfer/13-2009/pdf/013-016.pdf>>.
  5. PEKÁR, J. a kol., 2017: Modelovanie rozmiestňovania autorizovaných spracovateľov starých vozidiel na území Slovenskej republiky. In Konkurencia : proceedings of 9th annual international scientific conference : Jihlava.

#### **Kontaktné údaje**

prof. Mgr. Juraj Pekár, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 828

email: [pekar@euba.sk](mailto:pekar@euba.sk)

# ANALÝZA OBALU DÁT AKO NÁSTROJ VIACKRITERIÁLNEHO HODNOTENIA VARIANTOV

## DEA AS A MULTI-CRITERIA EVALUATION TOOL

*Michal Pieter*

### **Abstrakt**

Modely analýzy obalu dát (DEA) slúžia na hodnotenie jednotiek súboru, pre svoje vlastnosti však bývajú často využívané aj v teórii rozhodovania, a to väčšinou v kombinácii s niektorou z metód viackriteriálneho hodnotenia variantov (VHV). Tento príspevok sa zameriava na paralely medzi oboma oblasťami a skúma možnosti využitia modelov analýzy obalu dát samotných ako rozhodovacieho nástroja. V stručnosti predstavené sú tiež vybrané existujúce prístupy, vychádzajúce zo základných modelov DEA. Pokračovaním tohto príspevku sú plánované publikácie zameriavajúce sa na využitie pokročilých modelov a ich využitie v praxi.

*Kľúčové slová:* analýza obalu dát, DEA, viackriteriálne hodnotenie variantov, VHV

### **Abstract**

Data envelopment analysis (DEA) models are used to evaluate units within a group, however, due to their unique properties, they are often used in decision theory as well, usually in combination with one of the methods of multi-criteria evaluation (MCE). This paper takes a look at parallels between both fields and examines the possibility of using DEA models in and of themselves as a tool for decision support. Several existing approaches, utilizing some of the more basic DEA models, are surveyed and briefly explained. Upcoming planned publications will deal with the more advanced models, as well as their practical applications.

*Keywords:* data envelopment analysis, DEA, multi-criteria evaluation, MCE

## **1 ÚVOD**

Analýza obalu dát (data envelopment analysis, DEA) predstavuje súbor metód a nástrojov, ktoré umožňujú určiť relatívnu efektívnosť skúmaných jednotiek daného súboru, charakterizovanými určitými vstupmi a výstupmi. Viackriteriálne hodnotenie variantov (VHV, multi-criteria evaluation, MCE) je oblasť teórie rozhodovania poskytujúca metódy na určenie vhodnej voľby z diskkrétnej množiny možností na základe individuálnych preferencií rozhodovateľa. Medzi oboma prístupmi, resp. medzi DEA a viackriteriálnym rozhodovaním (multi-criteria decision analysis, MCDA, tiež MCDM) obecné, existuje viacero paralel, ktoré majú za následok ich aplikáciu v podobných situáciách. Viacerí autori sa obsiahlejšie venovali porovnaniu oboch prístupov, obzvlášť Belton a Stewart [1], ktorí po ich porovnaní dospeli k názoru, že predstavujú viacej doplnujúce sa, než protichodné techniky. Ostatne najmä v posledných rokoch im dáva za pravdu predstavenie viacerých metód využívajúcich spojenie či prepojenie oboch prístupov za účelom odstránenia nedostatkov jedného z nich vhodnou aplikáciou toho druhého s ohľadom na špecifiká konkrétnej aplikácie. Ako príklad môže poslúžiť napr. prepojenie DEA s metódou AHP [2] alebo metódou MACBETH [3]. Prepojenie oboch prístupov však nie je témou tejto práce. Tou je preskúmanie analýzy obalu dát ako priamej náhrady, t.j. v akých situáciách a za akých podmienok je možné pristupovať

k DEA ako k nástroju konceptuálne porovnateľnému s viackriteriálnym hodnotením variantov. Tento príspevok sa pritom zameriava na jednoduchšie formy DEA – využitie pokročilejších modelov, ako aj ich aplikácia v praxi budú témou pripravovaných prác.

## 2 TEORETICKÉ ASPEKTY

Pred predstavením konkrétnych možností využitia DEA ako nástroja pre podporu rozhodovania je vhodné oba prístupy porovnať a identifikovať paralely, ako aj potenciálne prekážky. Základnou otázkou, na ktorú je nutné odpovedať, je aké ciele sú očakávané ako od analýzy obalu dát, tak od viackriteriálneho hodnotenia variantov. Obidva prístupy sa v tradičnom ponímaní používajú v rozdielnych situáciách a očakávajú sa od nich rozdielne výsledky. V prípade DEA sa jedná o situácie, kedy je skúmaná daná dopredu určená množina, resp. súbor jednotiek, obvykle značného rozsahu. Cieľom je ohodnotiť každú z jednotiek na základe existujúcich a známych hodnôt zvolených metrík, vystupujúcich v modeli ako vstupy a výstupy (obecne faktory), pričom počet skúmaných jednotiek by ich mal výrazne prevyšovať. O DEA sa často hovorí ako o ex-post analýze, vzhľadom na to, že všetky údaje sú dopredu dané a zadávateľ analýzy do nej aktívne nevstupuje, je len konzumentom jej výsledkov. V stručnosti je možné povedať, že sa jedná o proces dodatočného, resp. retrospektívneho vyhodnotenia.

Viackriteriálne hodnotenie variantov ako oblasť teórie rozhodovania si kladie za cieľ z množiny alternatív určiť nejakú jej obmedzenú podmnožinu, ideálne jedinú alternatívu, ako v určitom zmysle slova najvhodnejšiu. Definícia vhodnosti pritom závisí na voľbe konkrétnej metódy, posudzovaných kritérií a preferencií rozhodovateľa, ktorý je spravidla aktívnym účastníkom rozhodovacieho procesu, či už v prípravnej alebo výpočtovej fáze. Podľa zvolenej metódy nemusia byť niektoré elementy ani vopred známe a konkrétne hodnoty sa zistia až samotnou aplikáciou. Nie je preto prekvapujúce, že VHV predstavuje vo valnej väčšine prípadov ex-ante analýzu, inými slovami sa jedná o techniku výberu a rozhodovania zameranú na budúcnosť.

### 2.1 DEA ako rozhodovací nástroj

Na prvý pohľad môže pôsobiť, že hlavnou prekážkou sú rozdielne definované ciele. Tú je však možné prekonať bližším preskúmaním výstupnej informácie DEA oproti metódam VHV. Metódy VHV je možné rozdeliť do troch základných skupín podľa charakteru tejto informácie:

- **Jediná alternatíva:** Metóda určí ako vhodnú jedinú alternatívu, vhodnosť ostatných alternatív nemusí byť ďalej posudzovaná, mimo fakt že nespĺňajú požiadavky.
- **Skupina alternatív:** Metóda určí ako rovnako vhodné určitú podmnožinu alternatív, diskrimináciu v rámci tejto podmnožiny ďalej neskúma.
- **Zoradenie alternatív:** Určenie poradia každej alternatívy podľa vhodnosti ich výberu.

V rozdelení podľa charakteru výstupnej informácie je možné badať pozoruhodnú podobnosť s modelmi DEA, či už tradičnými alebo ich modifikáciami. Tie prisudzujú každej jednotke mieru efektívnosti, v tradičných modeloch ako sú CCR [4], BCC [5] alebo SBM [6] od 0 do 1, kde 1 predstavujú efektívne jednotky. Na základe tejto kvantitatívnej miery je možné jednotky zoradiť a pri reinterpretácii jednotiek ako alternatív uvažovať tie s vyššou mierou efektívnosti ako vhodnejšie. Vzhľadom na to, že efektívnosť je v modeloch DEA relatívna k danému súboru, vždy je aspoň jedna jednotka určená ako efektívna, túto je potom možné považovať za vybranú.

Problém nastáva, keď je ako efektívnych určených príliš veľa jednotiek a analýza tak stráca svoj zmysel. K tomu často dochádza pri malom súbore, veľkému počtu faktorov (viď kapitolu 2.3), prípadne pri uvažovaní variabilných výnosov z rozsahu, ktoré Belton a Stewart považujú za vhodnejšie pre účely rozhodovania [1, s. 94]. Našťastie existujú rozšírené modely DEA, ktoré umožňujú úplne zoradenie jednotiek, napríklad modely super-efektivity ako je Toneho SSBM [7], kde miery presahujú 1. Tieto a ďalšie modely pre úplne zoradenie sú predstavené napríklad v [8] alebo [9]. V prípade, že počet efektívnych jednotiek nepresahuje prijateľnú hranicu je možné využiť aj tradičné modely v úlohe výberu skupiny zhodne vhodných alternatív. Tieto potom môžu byť detailnejšie preskúmané v druhej fáze rozhodovacieho procesu inou metódou. DEA tak môže poslúžiť na reštrikciu skúmaného súboru a elimináciu nevhodných alternatív. Zavedením obmedzení na prípustné váhy do modelu je pritom možné ovplyvniť počet efektívnych jednotiek. Viac o úlohe váh pojednáva kapitola 2.4.

## 2.2 Objektívnosť a subjektívnosť v DEA

Ako často prisudzovaná výhoda analýzy obalu dát býva citovaná jej objektívnosť. Je nepochybne pravda, že v porovnaní s metódami viackriteriálneho rozhodovania je menej ovplyvnená subjektívnymi názormi či už zadávateľa úlohy alebo analytika. Je však nutné si uvedomiť, že ani v DEA sa im nedá vyhnúť. Už samotná voľba modelu záleží na preferenciách koncového používateľa analýzy, hlavne čo sa týka charakteru požadovanej výstupnej informácie. Veľkým zdrojom subjektivity je voľba faktorov – vstupov a výstupov. Keď je k dispozícii väčšie množstvo faktorov s dostatočne kvalitnými dátami, výber niekoľkých z nich (s ohľadom na počet jednotiek v súbore), môže zásadným spôsobom ovplyvniť výsledok. Metódy teórie rozhodovania sú pritom založené na získaní a spracovaní subjektívnej informácie od rozhodovateľa. Tieto techniky môžu byť využité k riadenému spracovaniu subjektívnej informácie v DEA. Je pritom možné badať kompromis medzi objektivitou a využitím DEA ako rozhodovacieho nástroja. Napr. v kapitole 2.4 sa pojednáva o využití reštrikcie váh, na základe subjektívnych preferencií rozhodovateľa, k zvýšeniu rozlišovacej schopnosti modelu.

## 2.3 Faktory a kritériá

V metódach VHV sú alternatívy posudzované podľa zvolených kritérií. Sú to práve hodnoty kritérií pre danú alternatívu, ktoré primárne určujú vhodnosť jej výberu. V DEA sú miery efektivity zase závislé od hodnôt vstupov a výstupov pre každú jednotku, kolektívne označovaných ako faktory. Pri využití DEA na rozhodovanie môžu teda úlohu kritérií zastávať práve vhodne zvolené faktory. Minimalizačné kritériá je možné modelovať ako vstupy, keďže ich nižšia hodnota je pre jednotku výhodnejšia, a naopak maximalizačné kritériá sú predstavované výstupmi. Výhodou DEA je fakt, že získané hodnoty faktorov nie je nutné pred samotným výpočtom normalizovať na spoločnú škálu, ako je to v prípade viacerých metód VHV, napríklad TOPSIS, MAPPAC a iných. Rozdiel medzi DEA a VHV je možné badať v pomere počtu kritérií, resp. faktorov, k počtu alternatív, resp. hodnotených jednotiek. Metódy VHV obecné vyberajú z relatívne malého počtu alternatív na základe veľkého počtu kritérií. V DEA je takáto situácia problematická – vzhľadom na slobodnú voľbu váh faktorov, bližšie popísanú v nasledujúcej kapitole, bývajú pri nízkom počte jednotiek, v pomere k faktorom, mnohé z nich, prípadne až všetky, označené ako efektívne. Pre dostatočnú rozlišovaciu schopnosť sa preto zvykne uvádzať pravidlo zvoliť taký počet faktorov (vstupov a výstupov dohromady), aby počet jednotiek, ktorý je obyčajne daný, bol aspoň trojnásobný. [10, s. 92-93]



## 2.4 Váhy kritérií

Podstatnou vlastnosťou väčšiny modelov DEA je možnosť každej jednotky voliť váhy jednotlivých faktorov tak, aby výsledná miera efektívnosti bola čo najvyššia a ona tak určená ako čo možno najvhodnejšia. Samotné hodnoty týchto váh sú pritom určené optimalizáciou konkrétneho matematického modelu. V prípade absencie akejkoľvek subjektívnej informácie síce môže byť takéto arbitrárne určenie váh akceptovateľné, v reálnej praxi to však spravidla nebýva. Situácie, kedy je význam prisúdený malej skupine, či jedinému kritériu a ostatné majú nulové váhy, prípadne zhodná dôležitosť všetkých kritérií sú v skutočnom svete výnimkou, nie pravidlom. Existuje preto skupina modelov, ktoré umožňujú zachytiť preferencie rozhodovateľa vo forme obmedzení váh pre jednotlivé faktory, pričom v rámci takto daných mantinelov môže model určiť jednotke tie najvhodnejšie váhy. Tu je znovu možné badať už zmienený kompromis – síce sa zvýšila užitočnosť nástroja pre rozhodovanie odstránením pre rozhodovateľa neprijateľných možností, zároveň sa však znížila objektivita modelu. Dodanie subjektívnej informácie si vyžaduje väčšie zapojenie rozhodovateľa a tiež spracovanie jeho preferencií do vhodnej matematickej podoby, tak aby ich bolo možné využiť v modeli.

Určitým odklonom od tradičných modelov sú prístupy, ktoré prisudzujú tie isté váhy faktorov všetkým jednotkám súboru. Medzi často uvádzané výhody týchto modelov so spoločnými váhami (common-weights) býva zmyslupnnejšie zoradenie a menšia výpočtová náročnosť, keďže vo väčšine z nich sa optimalizuje len jeden matematický model pre celý súbor. Jednotlivé modely sa líšia spôsobom, akým tieto váhy stanovujú, čo do veľkej miery závisí od sledovaného cieľa, prípadne požadovania dodatočnej informácie v podobe preferencií. Obecné kategórie prístupov sú popísané napríklad v [11, s. 103-04]. Niektoré z modelov spoločných váh vystupujú aj v podobe rozhodovacích nástrojov a sú v stručnosti uvedené v kapitole 3.4.

## 3 VYBRANÉ PRÍSTUPY

Ako už bolo uvedené v úvode, existuje veľa techník využívajúcich analýzu obalu dát v kombinácii s nejakou metódou viackritériálneho hodnotenia variantov. Dočítať sa o nich je možné z viacerých už spomenutých zdrojov, témou tohto príspevku sú však prístupy a príklady využitia modelov DEA a ich modifikácií ako priamych nástrojov na podporu rozhodovania. Tých už je menší počet a niektoré z nich sú predstavené na nasledujúcich riadkoch.

### 3.1 Nemodifikované modely

Najjednoduchšou možnosťou je priama aplikácia niektorého z tradičných modelov na rozhodovací problém. Doyle a Green [12] zvolili CCR model a využili ho pri výbere nákupu produktov spotrebnej elektroniky. Ako kritériá poslúžili faktory, konkrétne 1 vstup reprezentujúci minimalizačné kritérium ceny a 7 výstupov pre vlastnosti produktu – všetky maximalizačného charakteru. Bez zadania dodatočných preferencií, napríklad vo forme obmedzenia váh sa u nich vyskytol problém načrtnutý v kapitole 2.1 – z 37 produktov ich bolo 11 určených ako efektívnych. Ako možnosť ich rozlíšiť navrhli zamerať sa na neefektívne produkty – pre každý z nich je možné skonštruovať virtuálnu jednotku, definovanú konvexnou lineárnou kombináciou niektorých referenčných efektívnych produktov. Tie tak predstavujú vzory pre zvýšenie efektívnosti neefektívnych jednotiek. Efektívne produkty sa preto rozhodli zoradiť zostupne podľa počtu ich výskytu ako vzorov pre tie neefektívne – produkt s najväčším výskytom bol nakoniec vybraný ako najvhodnejší.

Papagapiou a i. [13] potom využili rovnakú techniku pri voľbe osobných počítačov pomocou modelu BCC.

Bouyssou [10] upozorňuje na problémy spojené s využitím modelov DEA takýmto spôsobom, t.j. bez vhodnej úpravy s ohľadom na špecifiká rozhodovacích úloh a využitie subjektívnej informácie. Jedným z problémov je fakt, že napriek tomu že všetky jednotky určené ako efektívne sú nedominované (v zmysle teórie rozhodovania), naopak to však neplatí – nedominovaná jednotka, ktorá by mohla byť pre rozhodovateľa zaujímavá, nemusí byť nutne označená za efektívnu. Okrem iných upozorňuje aj na opačný problém v prípade, že je snaha zoradiť jednotky ako vhodné podľa klesajúcej miery efektívnosti. V prípade, že efektívna jednotka A dominuje jednotku B, ktorá má len nepatrne nižšie hodnoty všetkých kritérií, B nemôže byť označená ako efektívna a bude určená ako menej vhodná než všetky efektívne jednotky. To platí aj v situácii, ak by B ďaleko prekonávala ostatné jednotky C, D, atď. podľa všetkých kritérií okrem jedného, v ktorom by mierne zaostávala. V extrémnom prípade sa môže stať, že by bola označená ako najmenej vhodná, a to napriek tomu, že každý racionálny rozhodovateľ by ju označil za druhú najvhodnejšiu. Ten istý problém sa môže vyskytnúť aj pri použití modelov super-efektivity, kde navyše poukazuje na náchylnosť zmeny poradia efektívnych jednotiek na zaradenie, resp. vyradenie jednotiek zo súboru. [10, s. 975-76]

### 3.2 Modely s obmedzenými váhami

V záujme zlepšenia využitia modelov DEA ako praktického rozhodovacieho nástroja sa viacerí autori rozhodli zakomponovať preferencie rozhodovateľa do modelu vo forme obmedzení váh, ako to bolo popísané v kapitole 2.4. Sarrico a i. [14] využili tento prístup pri použití DEA na rozhodovanie študentov pri výbere univerzity v Spojenom kráľovstve. Uvedomili si nielen to, že študenti prikladajú jednotlivým z 10 určených kritérií iný význam, ale aj to, že medzi záujemcami o štúdium sa môžu preferencie výrazne odlišovať. Určili preto 6 typových študentov a riešili úlohu pre každého z nich samostatne. Obmedzenia váh získali tým, že roztriedili kritériá, resp. faktory, do troch kategórií podľa klesajúceho významu pre daného typového záujemcu. Podľa počtu kritérií v jednotlivých kategóriách určili na základe jednoduchého vzorca spodné a horné limity pre váhy zodpovedajúcich faktorov tak, aby významnejšie dostali vždy väčšie váhy ako tie menej významné, a zároveň, aby neboli žiadne faktory opomenuté. Tieto obmedzenia nakoniec využili v tradičnom CCR modeli k získaniu mier efektívnosti a zoradeniu univerzít podľa nich ako mier vhodnosti výberu. Mimochodom, práve modely s obmedzenými váhami bývajú často prepájané s vybranými metódami teórie rozhodovania, príkladom môže byť už spomenuté prepojenie s metódou MACBETH [3].

### 3.3 Krížová efektívnosť

V kapitole 3.1 bolo spomenuté využitie modelov super-efektivity na zoradenie jednotiek a tým určenie vhodnosti ich výberu. Tie však nepredstavujú jediný spôsob zoradenia celého súboru jednotiek, jednými z ďalších sú modely krížovej efektivity, resp. hodnotenia (cross-efficiency, resp. cross-evaluation). V stručnosti, tieto sú založené na aplikácii optimálnych váh skúmanej jednotky na ostatné jednotky súboru a tým určenie ich mier efektívnosti, ak by využívali rovnakú štruktúru ako jednotka skúmaná. Vypočítaním týchto mier pre všetky skúmané jednotky sa získa matica krížových efektívít rozmeru  $n \times n$ , kde  $n$  je počet jednotiek súboru.

Doyle [15] navrhol metódu AXE aplikujúcu koncept krížovej efektivity v DEA na rozhodovacie problémy, konkrétne výber univerzity. Pri určení váh skúmaných jednotiek spravidla existuje viacero optimálnych riešení v istom rozsahu, preto Doyle zavádza na ich obmedzenie sekundárny cieľ minimalizácie krížových efektívít ostatných jednotiek, tzv. agresívny prístup. Každá skúmaná jednotka preto volí váhy tak, aby bola nielen čo

najefektívnejšia, ale aj najvýnimočnejšia. Následne pre každý faktor–kritérium určí priemernú hodnotu týchto váh. Takto získaný vektor váh interpretuje ako preferencie priemerného rozhodovateľa. Ako mieru vhodnosti, podľa ktorej je možné jednotky zoradiť, volí nakoniec súčet hodnôt faktorov pre každú jednotku, vážený takto získanými priemernými váhami kritérií. Jeho metóda tak nevyžaduje zadanie preferencií rozhodovateľom, ale vychádza z preferencií modelovaného fiktívneho, priemerného, rozhodovateľa.

Mavrotas a Trifillis [16] navrhli rozšírenie metódy AXE, nazvané ACE+. Táto umožňuje každej jednotky zvoliť si nielen váhy, ale aj charakter funkcie úžitku tak, aby boli pre ňu čo najvhodnejšie. Tvar monotónne rastúcej funkcie úžitku je v modeli zahrnutý pomocou premennej, ktorej hodnota určuje do akej miery je konvexná či konkávna, prípadne lineárna. ACE+ taktiež uvažuje odlišnú účelovú funkciu, kde namiesto primárneho a sekundárneho cieľa maximalizuje rovno rozdiel medzi mierou efektívnosti skúmanej jednotky a priemerom efektívity ostatných, čím jej dáva najväčšie šance pri finálnom hodnotení. Toto hodnotenie je zistené z matice krížových efektívity totožným spôsobom ako v metóde AXE.

Využitie krížovej efektivity odstraňuje problém, kedy všetky neefektívne jednotky sú určené ako menej vhodné než všetky efektívne, ako tomu bolo v modeloch super-efektivity. Bouyssou však upozorňuje na komplikovanú závislosť hodnotenia jednotky na ostatných členoch súboru. Problém vidí v prípadnej chýbajúcej monotónnosti, kedy zhoršenie hodnôt jednotky môže mať za následok zlepšenie vhodnosti jej výberu oproti niektorým iným. [10, s. 977]

### 3.4 Ďalšie prístupy

Existujú aj ďalšie modely využívajúce DEA ako nástroj pre podporu rozhodovania. Jedným z nich je CWDEA Torabiho a Shokra [17], ktorý patrí medzi modely so spoločnými váhami a je schopný určiť miery efektívnosti všetkých jednotiek naraz. Jeho zvláštnosťou je špeciálne zachytenie hodnôt kvalitatívnych faktorov – v podobe kategoriálnych dát – v modeli, oddelene od faktorov kvantitatívnych, čím odpadá potreba ich najprv transformovať. Model aplikujú na výber vhodného materiálu pre konštrukciu zotrvačiek v priemysle. Nevýhodou je, že nezaručuje jednoznačné zoradenie všetkých jednotiek. Autori neskôr skonštruovali verziu s vylepšenou schopnosťou diskriminácie, ktorá tento problém zmiernuje [18].

Darehmiraki a Behdani [19] navrhli zaujímavý model, ktorý nie je založený na mierach efektívnosti samotných jednotiek, ale na efektívnosti fiktívnej agregovanej jednotky, ktorej hodnoty kritérií sú dané ako súčet hodnôt všetkých jednotiek. Modelom CCR najprv vypočítajú jej mieru efektívnosti. Následne pre každú jednotku súboru skúmajú, ako sa zmení efektivita tej agregátnej, ak ju z nej v ďalšej úvahe vynechajú. Rozdiel medzi mierou efektívnosti agregátnej jednotky pred a po vynechaní skúmanej jednotky vyjadruje mieru jej príspevku k efektivite agregátnej jednotky, pričom vyššia hodnota je považovaná za lepšiu. Na základe tejto miery sú nakoniec všetky jednotky zoradené. Model tak umožňuje úplne zoradenie, autori však upozorňujú, že je vhodný len pre rozhodovacie problémy s maximalizačnými kritériami.

## 4 ZÁVER

V priebehu tohto príspevku boli predstavené základné teoretické aspekty a možnosti využitia analýzy obalu dát v úlohe nástroja pre podporu rozhodovania, špeciálne vo forme viackritériálneho hodnotenia variantov. Následne bolo stručne predstavených niekoľko existujúcich prístupov, vychádzajúcich zo základných modelov DEA. Prehľadom podobného využitia pokročilejších modelov DEA v komplikovanejších rozhodovacích situáciách sa bude zaoberať pripravovaná ďalšia publikácia. Špeciálne zameranie by pritom malo byť venované

modelom s nepresnými informáciami (fuzzy DEA) a vnútornou štruktúrou jednotiek (siet'ová, network DEA). Nakoniec, zavŕšením tejto série by malo byť praktické využitie predstavených modelov v pripravovanom aplikačnom článku v kolaborácii s ďalšími autormi.

### Použitá literatúra

1. BELTON, V. a T. J. STEWART. DEA and MCDA: Competing or Complementary Approaches? In: *Advances in Decision Analysis*. Springer Netherlands, 1999, s. 87-104. ISBN: 978-90-481-5167-7 978-94-017-0647-6.
2. SINUANY-STERN, Z., A. MEHREZ a Y. HADAD. An AHP/DEA methodology for ranking decision making units. 2000, **7** (2), 109-24. ISSN: 0969-6016.
3. VIEIRA JUNIOR, H. Multicriteria approach to data envelopment analysis. 2008, **28** (2), 231-42. ISSN: 0101-7438.
4. CHARNES, A., W. W. COOPER a E. RHODES. Measuring the efficiency of decision making units. 1978, **2** (6), 429-44. ISSN: 0377-2217.
5. BANKER, R. D., A. CHARNES a W. W. COOPER. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. 1984, **30** (9), 1078-92. ISSN: 0025-1909.
6. TONE, K. A slacks-based measure of efficiency in data envelopment analysis. 2001, **130** (3), 498-509.
7. TONE, K. A slacks-based measure of super-efficiency in data envelopment analysis. 2002, **143** (1), 32-41. ISSN: 0377-2217.
8. ADLER, N., L. FRIEDMAN a Z. SINUANY-STERN. Review of ranking methods in the data envelopment analysis context. 2002, **140** (2), 249-65. ISSN: 0377-2217.
9. HOSSEINZADEH LOTFI, F. a i. A Review of Ranking Models in Data Envelopment Analysis. *Journal of Applied Mathematics*, 2013.
10. BOUYSSOU, D. Using DEA as a tool for MCDM: some remarks. 1999, **50** (9), 974-78. ISSN: 0160-5682, 1476-9360.
11. ROLL, Y., W. D. COOK a B. GOLANY. Controlling Factor Weights in Data Envelopment Analysis. 1991, **23** (1), 2-9. ISSN: 0740-817X.
12. DOYLE, J. a R. GREEN. Comparing products using data envelopment analysis. 1991, **19** (6), 631-38. ISSN: 0305-0483.
13. PAPAGAPIOU, A., J. MINGERS a E. THANASSOULIS. Would you buy a used car with DEA. 1997, **10** (1), 13-19. ISSN: 1759-0477.
14. SARRICO, C. S. a i. Data Envelopment Analysis and University Selection. 1997, **48** (12), 1163-77. ISSN: 0160-5682.
15. DOYLE, J. R. Multiattribute Choice for the Lazy Decision Maker: Let the Alternatives Decide. 1995, **62** (1), 87-100. ISSN: 0749-5978.
16. MAVROTAS, G. a P. TRIFILLIS. Multicriteria decision analysis with minimum information: combining DEA with MAVT. 2006, **33** (8), 2083-98. ISSN: 0305-0548.
17. TORABI, S. A. a I. SHOKR. A Common Weight Data Envelopment Analysis Approach for Material Selection. 2015 roč. XXVIII, s. 913-21.
18. SHOKR, I., M. SADEGH AMALNICK a S. A. TORABI. An Augmented Common Weight Data Envelopment Analysis for Material Selection in High-tech Industries. 2016, **3** (2), 1234-52.
19. DAREHMIRAKI, M. a Z. BEHDANI. A new DEA approach to rank alternatives in MCDA. 2013 roč. MMXIII. ISSN: 21954496.

**Kontaktné údaje**

Ing. Michal Pieter

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky

Nám. W. Churchilla 1938/4, 130 67 Praha 3

Tel: (+420) 224 09 5445

email: [michal.pieter@vse.cz](mailto:michal.pieter@vse.cz)

# MOSAIKOVÉ GRAFY JAKO NÁSTROJ PRO VIZUALIZACI DAT A TVORBU MODELU: PŘÍPADOVÁ STUDIE PRO ŠPANĚLSKÝ PRACOVNÍ TRH

## MOSAIC PLOTS AS A TOOL FOR DATA VISUALISATION AND MODEL CREATION: A CASE STUDY FOR SPANISH LABOUR MARKET

*Tatiana Polonyankina*

### Abstrakt

Mosaikové grafy jsou analytickým nástrojem používaným k vizualizaci a popisu kategoriálních dat. Technika umožňuje jednoduchým způsobem prezentovat výsledky analýzy a zastoupení jednotlivých kategorií, přesto je však zřídka používána v ekonometrických a statistických člancích. Hlavní výhodou je schopnost pojmout velké množství informací v jenom grafu, kdy se kombinuje popis dat pomocí ploch v grafu a zobrazení reziduí odhadnutého log-lineárního modelu. Příspěvek v první části popisuje techniku tvorby grafů a ve druhé části aplikuje metodu na Španělská roční individuální data z průzkumu pracovních sil Eurostat 2014 s cílem najít a popsat rozdíly mezi cizinci a Španěly.

**Klíčová slova:** *mosaikový graf, logaritmicko-lineární modely, migrace, Španělsko*

### Abstract

Mosaic plotting is a data-analytic tool used to visualise basic characteristics of a categorical data. The technique gives an easy understandable way of a results presenting however is rarely used in econometric and statistic papers. The main advantage is an ability to cover large amount of information in one plot and possibility to combine visual proportional description techniques with model error deviation patterns. The paper describes the method of mosaic plots creation and applies the method on annual data from Spanish Labour Force survey, concentrating on searching for patters. The log-linear models of complete and joint independence are used for plotted residuals in mosaic plots.

**Keywords:** *mosaic plots, log linear models, immigration, Spain*

## 1 INTRODUCTION

The mosaic plotting method is applied on contingency tables where categories are displayed with respect to their size and proportion in a sample. Main advantage is an ability to analyse how a dataset fits the baseline model of independence or partial independence and use colour-shading to visualise residuals (departures from the model). The colour-shading is used to represent a sign and a magnitude of standardized residuals from a specified model. The mosaic display in its basic way was introduced by Hartigan and Kleiner [2], [3] where the cell area is proportional to cell frequency showing that visual comparison of cells helps to highlight large and small counts in data set. The paper written by Friendly [1] is summarising a use of contingency tables giving a few examples as an application demonstration. The visualisation of data summarised in a contingency table and residuals gives the opportunity to detect and highlight some patterns as well helps to construct a model to describe dependencies between variables.

The classical log-linear models, as tools for analysing relationships among two or more variables where a response variable is categorical, are estimated within a mosaic plot. In case of application of the models on contingency tables there are these specifications: the response and explanatory variables are all categorical, the observations consist of frequencies in the

cells of a contingency table. Next standardised residuals are calculated and displayed by shading in the mosaic. Since the method is working with contingency tables, the main application field is probably medicine, this article is trying to connect the method with analysis of immigration and Spanish labour market between 2011 and 2014. Spain was chosen since it is a country with one of the highest immigration as well as country with higher unemployment rate in European union during the analysing time period (more is described in chapter 3).

## 2 MOSAIC PLOTS FOR MULTI-WAY CONTINGENCY TABLES

Multi-way tables<sup>1</sup> are data sets with more than two groups of variables, where a contingency table is defining multi-way relationships between frequencies (cells). Mosaic plotting gives information about joint distribution at several levels:

- cell frequencies are still represented by the area of a tile,
- a spacing scheme allows visually combine cells for marginal relations between cells,
- shading with respect to residuals to show fitting of a model.

Let's suppose that  $N$  is a table with frequencies of variables groups A, B and C, where  $i = 1, \dots, m; j = 1, \dots, n$  and  $k = 1, \dots, o$ . The contingency table and its values are defined as:  $N = \{n_{ijk}\}$ . The log-linear complete independence model for the multi-way contingency table defined above is:

$$n_{ijk} = e^{\mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}}, \quad (1)$$

$i = 1, \dots, m; j = 1, \dots, n$  and  $k = 1, \dots, o$ . Under constrains:  $\sum_i \alpha_i = 0, \sum_j \beta_j = 0, \sum_k \gamma_k = 0$ ,  $n_{ijk}$  is observed frequency for variable  $A_i$  from variable group A,  $B_j$  from variable group B and  $C_k$  from variable group C,  $\mu$  is aspect common to all frequencies in the table,  $\alpha_i$  is the effect of variables from group A,  $\beta_j$  is the effect of variables from group B,  $\gamma_k$  is the effect of variables from group C,  $\varepsilon_{ijk}$  is a random error term.

The parameters  $\mu, \alpha_i, \beta_j, \gamma_k$  are estimated using Maximal likelihood method. Expected frequencies are derived from the equation  $\hat{m}_{ijk} = e^{q_{ijk}}$ , the value  $q_{ijk}$  is obtained from estimated parameters of the log-linear independence model defined above,  $q_{ijk} = \log(\hat{n}_{ijk})$ .

The primary goal of the mosaic plot is to describe the data, not to find the perfect model. However, this article is showing how to use mosaic plots to find a log-linear model that fits a contingency table data. When the displayed standard errors for (1) are high the other possibility is to use a log-linear model with partial independence defined in (2). If a partial independence model still displays high residuals, the higher-order model of log-linear models group can be used.

$$n_{ijk} = e^{\mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta_{ij}) + (\beta\gamma_{jk}) + (\alpha\gamma_{ik}) + \varepsilon_{ijk}}, \quad (2)$$

For easier interpretation and visualisation residuals displayed in mosaics are standardised. There are several possibilities to standardise residuals. R gives the options of Pearson residuals, standardisation based on components of the likelihood or the Freeman-Tukey residuals. In majority of literature as well as in this text Pearson residuals are used. A standardized residual for cell  $i, j$  is for Pearson  $X^2$  are done following the equation:

---

<sup>1</sup> Mosaic plots can be used as well with two-way contingency tables which is not included in the article, for more information and example see [1].

$$d_{ijk} = \frac{(y_{ijk} - \hat{m}_{ijk})}{\sqrt{\hat{m}_{ijk}}}.$$

Residuals  $d_{ij}$  as a measure of standardized deviation from independence are visualized by the shading of each rectangle. Positive deviation is displayed with a solid line and negative deviation with a broken line. The pigment pattern describes  $d_{ij}$  distribution where: cells with  $|d_{ij}| \leq 2$  are empty, cells with  $|d_{ij}| \geq 2$  and  $|d_{ij}| \leq 4$  are filled, cells with  $|d_{ij}| \geq 4$  are filled with a darker colour. When  $|d_{ij}| \geq 4$  and most of the colours on a mosaic plot are darker used model does not fit a joint/complete independence model a researcher should think about changing the used model. Results from a mosaic plot are compared with 2 statistical tests in this article: The independence between more than two variables is tested by Chi-squared test where  $X^2$  is calculated using equation (3) with decrease of freedom  $df = (mno - 1) - (m + n + o - 3)$ . The null hypothesis of independence for all factors is tested.

$$X^2 = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^o \frac{(y_{ijk} - \hat{m}_{ijk})^2}{\hat{m}_{ijk}}, \quad (3)$$

Choice of joint independence model is added by a significant test of new included parameters. Specific parameters are tested using equation:  $\Delta X^2 = X_a^2 - X_b^2$ ,  $\Delta df = df_a - df_b$  where  $\Delta X^2$  is difference between  $X_a^2$  of a model without a tested parameter and  $X_b^2$  a model including a tested parameter. When  $\Delta X^2$  is significant, it suggests that the parameter is significant.

Basic structure of a mosaic plot has high and width of each box proportional to the marginal total and the area is proportional to expected frequency. In this article Condensed Column Proportion Mosaic, what is a more common way to display data where each rectangle shows the conditional frequency of rows given columns. The use of Condensed Column Proportion Mosaic can highlight some patterns but rows are harder aligned. With this graph, it is easier to compare columns but harder to compare rows, that means that researcher should think about variables and their placement in the graph.

### 3 A CASE STUDY: SPANISH LABOUR MARKET

Spanish labor market was chosen thanks to its specific position between European countries between 2011 and 2014: Spain was suffering from crisis where overall unemployment was one of the highest in Europe, Spain belong to countries with a high number immigrants living in the country. Combination of these factors gives an unique chance to examine differences between unemployment for immigrants and natives. This chapter is divided on three parts: First is summarizing Spanish migration and labor market, second describes the data used and the last part is searching for patterns in Spanish labor market using mosaic plots and log-linear regressions.

#### 3.1 Migration and Labour Characteristics for Spain

De la Rica and col. [5] summarized main empirical studies and described immigration trends in Europe. Spain has experienced increase of immigration over the last few decades, with particularly fast growth rates in Spain since 2000. The dramatic increase started in 2000 with 4 % of immigrants growing to share over 14 % in 2011<sup>2</sup>. The largest immigration group in Spain are Romanian, followed by Moroccan and Ecuadorian. Large group of working foreign-born population is formed by Latin Americans. Linguistic similarity is one of important immigra-

---

<sup>2</sup> For more details and data sources see [5].



tion push factor for Romanians and Latin Americans. Looking at Eurostat official statistics<sup>3</sup> number of persons born abroad, whether in other EU Member States or non-EU countries, was approximately 6 million between 2011 and 2014, that is ten times more than European median (560 thousands). These numbers put Spain between 5 European countries with higher number of foreign born population. This group includes Germany, United Kingdom, France and Italy. Unemployment rate in Spain in 2004 was 26.2 % [10]. Spanish labor market is suffering by two characteristic problems: Young people are worst affected by difficulties in accessing the labor market with 55.5 % unemployment rate in 2014 [10]. As a result, young people have been often choosing to prolong their education, enter a precarious short-term labor market or emigrate. Long-term unemployment was 13% in 2014 and represents another structural problem. Several steps were done to heal Spanish labor market [10] as government austerity policies aligned with Brussels directives, the 2012 labor reform introduced flexibility to a traditionally rigid job market.

### 3.2 Data

Two datasets are used: First is Eurostat official database used in section 3.1 and second individual data from European Labour Force Survey (LFS) are used to create contingency tables for mosaic plots and log-linear regressions. Only economically active population is analysed: individuals between 15 and 66 are included, inactive individuals and individuals in compulsory military service are excluded. Immigrant is an individual with country of birth different to the country of stay at a survey year. Recent immigrants are defined as individuals with year of residence in a country of stay lower or equal to 5. Settled immigrants are those with residence higher than 5 years in the country that is different to the birth country.

Table below shows distribution of individuals in LFS from 2011 to 2014. Data includes 50 thousand individuals each year with unemployment share between 20% - 26%, which approximately corresponds with official unemployment numbers. Foreign share in economically active population is about 7%, which shows that this group is underrepresented in the used sample comparing to total share of immigrants in Spain.

	2011	2012	2013	2014
Natives	47 191	49 209	47 655	47 679
Foreign - born	3 432	3 443	3 174	3 055
TOTAL	50 623	52 652	50 829	50 734
Unemployed	10 124	12 668	12 825	12 346
Employed	40 499	39 984	38 004	38 388

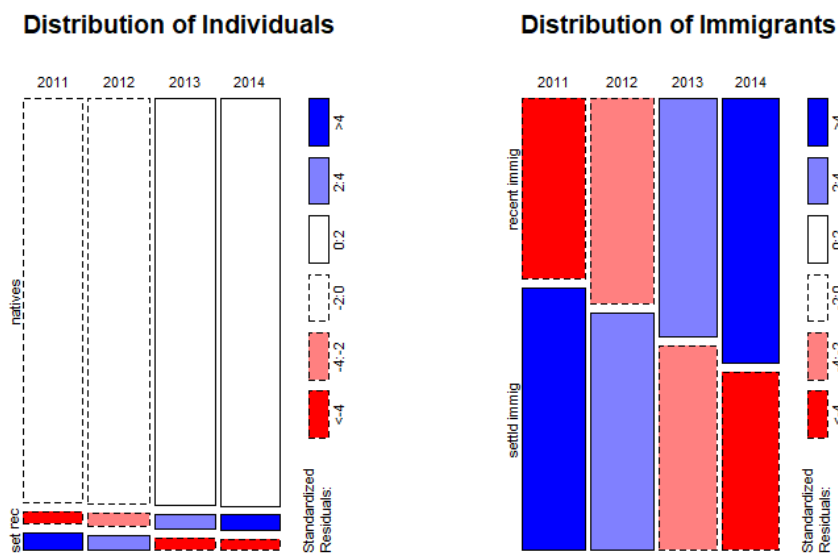
**Table 1** Labor Force Survey Data

### 3.3 Mosaic Plots and Log-Linear models

The left mosaic plot on Figure 2 shows distribution of individuals in the data set for each year. Number of economically active natives seems to be stable over time with a slight increase from 2013. The share of natives is significantly higher than number of individuals (around 93% of active population). Dark colours for immigrants on the first mosaic are zoomed on the right mosaic plot. The mosaics show that there are visible changes in immigrants over time where share of recent immigrants is increasing significantly and share of settled immigrants is decreasing. This may suggest that immigration to Spain continues even though the economic crisis hit Spain and stopped GDP increase. On the other hand settled immigrants seem to leave the economically active population which may be caused by retirement or emigration from Spain. Settled immigrants know the European labour market and the European culture,

<sup>3</sup> Numbers were taken from Eurostat database downloaded 27.4.2017 (<http://ec.europa.eu/eurostat/data/database>)

they also tend to be more experienced workers. Since the Economic crisis increased unemployment rapidly in Spain, this group may decide to leave the country and search for better opportunities abroad.



**Figure 2** Distribution of Individuals in the Analysis

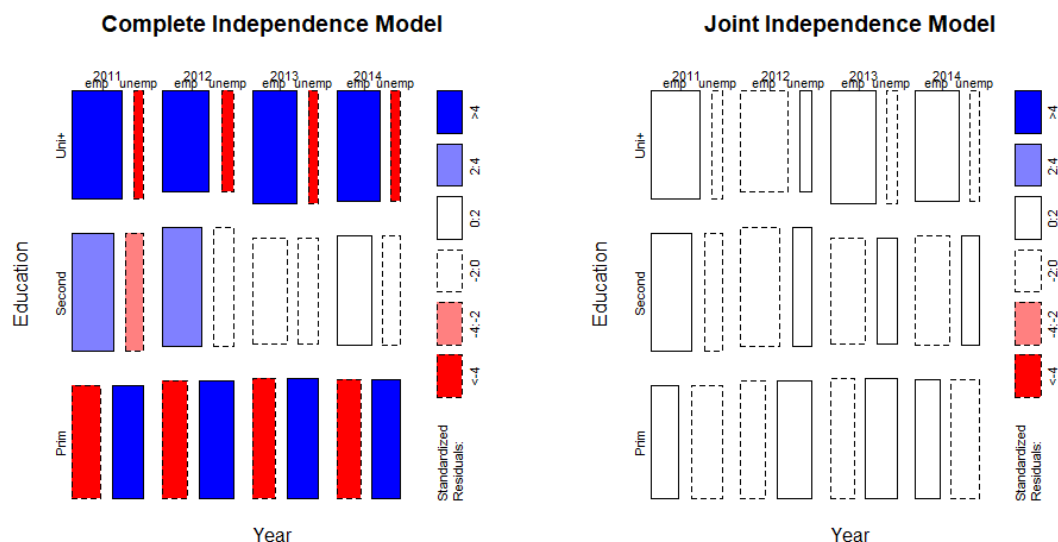
The second application (Figure 3) is looking closer on immigrants in Spain grouping them together with respect to their education and labour status between 2011-2014. Education has three common used groups: primary, secondary and university. Labour status is employed or unemployed at moment of the survey collection. The log-linear model of complete independence (1) is used in the left mosaic plot.

The main visible patterns are: The education of immigrants seems to not change much over time. The most of migrants have secondary and lower education. Unemployment rate has an expected evolution with respect to education level: immigrants with university and higher education tend to have lower unemployment rate than the other two groups. The highest share of unemployed immigrants seems to be in 2012. Looking on residuals opposite pattern is visible for university and primary educated immigrants: probability to be employed is much higher than average for university education for all years and as opposite probability to be unemployed is significantly higher for immigrants with primary education for all years.

Chi-squared test of complete independence between education, employment and year of survey was done. The null hypothesis of variables independence was rejected on 1% significance level. The same is shown looking at displayed residuals, where absolute values of standardised residuals are mostly higher than 4. The right mosaic plots residuals of the joint independence model (2) applied on data, where dependence for labour status and year is not included due to rejection of hypothesis of parameters significance. Mosaic plot shows that model of joint independence fits the contingency table data and the absolute values of standardised residuals are lower than 2 for all frequencies.

To make analysis complex the same model of joint independence (2) excluding dependence within labour status and year is done for natives. Both models were estimated using maximum likelihood method. Table 2 compares estimated coefficients for a joint dependence between labour status and education for immigrants and natives, where negative values are displayed bold. It is visible that education plays more vital role in employment status for immigrants than natives, since all coefficients in absolute values are higher for immigrants than na-

tives. The second visible pattern is that for immigrants it is “enough” to finish secondary education to improve probability to be improved, however for natives increase of probability to be employment is visible from university education. This can be related to the facts that share of immigrants with primary education is higher than share of natives and immigrants in average tend to work in lower/less professional positions than natives. D’Amuri and Peri (2011) shows that countries immigration has led native workers to specialize in occupations involving relatively complex tasks, that can relate to higher education requirement. Authors is [5] conclude that there are large gaps in labour market outcomes between immigrants and natives in most countries, both in terms of employment and wages. This can be also with education required for job positions.



**Figure 3** Mosaic Plots of Immigrants by Education and Employment Status

	Immigrants		Natives	
	Employed	Unemployed	Employed	Unemployed
University	0.42	<b>-0.42</b>	0.25	<b>-0.25</b>
Second	0.02	<b>-0.02</b>	<b>-0.00</b>	0.00
Primary	<b>-0.44</b>	0.44	<b>-0.25</b>	0.25

**Table 2** Comparison of Immigrants vs. Natives

## 4 DISCUSSION AND CONCLUSIONS

The article is introducing mosaic plots as a method to display contingency table data, that gives the opportunity to search for patterns as well as a tool to examine the log-linear model fitting. Log-linear models of complete and joint independence are used and parameters are estimated using maximal likelihood method. Model fitting and parameter significance is as well tested by Ch-squared test.

A case study of Spanish labour market is done in the application part. Given that Spain belong to the countries with high immigration and suffered for high unemployment between 2011 and 2014, the article concentrates on describing differences in employment patterns between natives and immigrants. Using mosaic plots it is shown that share of recent immigrants was increasing and opposite share of recent immigrants was decreasing in time period of analysis. Mo-

saic plots are used to demonstrate that a complete independence model does not fit the data and that a joint independence model significantly reduced residuals and better describes the data. Finally, regressed coefficients of joint probability of education and employment for immigrants were compared for parameters of the same model for natives, showing that dependence is higher for immigrants. Probability to be employed increase from secondary education for immigrants and from university education for natives.

Next steps that naturally comes to discuss and examine are an impact of immigration on Spanish labour market. Empirical studies [6], [7], [8] conclude that immigration does not reduce the employment rates of natives, and found no evidence of effect of immigration on native wages. More over native workers shifting from less to more complex occupations and [9] shows that female immigration to Spain has increased the local availability of household services and reduced their prices.

### Acknowledgements

The work on this paper was supported by the grant IGS F4/93/2017 of University of Economics, Prague.

### References

- [1] Friendly, M: Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, 89 (1994), 190–200.
- [2] Hartigan, J.A., and Kleiner, B.: A mosaic of television ratings. *The American Statistician*, New York: Springer-Verlag, 38 (1984) , 268-273.
- [3] Hartigan, J.A., and Kleiner, B.: Mosaics for Contingency tables. *Computer science and statistics; Proceeding of the 13th Symposium on the Interface*, 38 (1981), 32–35.
- [4] Dobson, J.A.: *An Introduction to Generalized Linear Models*, London: Chapman and Hall, (1990) ISBN 0412311003
- [5] De la Rica, S., Glitz, A., Ortega, F.: *Immigration in Europe: Trends, Policies and Empirical Evidence*, IZA Discussion Paper No. 7778 (2013), Bonn: Institute for the Study of Labour.
- [6] D’Amuri, Francesco and G. Peri, *Immigration, Jobs and Employment Protection: Evidence from Europe*, NBER Working Paper 17139, (2011).
- [7] De la Rica, Sara and T. Polonyankina, *The impact of Immigration on Occupational Specialisation among Natives in Spain: Does the Business Cycle Matter?* Forthcoming in *Revista de Economía Aplicada*, 63 (2013), 51 -74.
- [8] Carrasco, Raquel, J. F. Jimeno, and A. C. Ortega, *The Effect of Immigration on the Labor Market Performance of Native-born Workers: Some Evidence for Spain*, *Journal of Population Economics* (2008), Vol. 21 (3), 627-648.
- [9] Farré, L., L. González and F. Ortega (2011), "Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women", *The B.E. Journal of Economic Analysis & Policy*, Vol. 11 (1), Article 34.
- [10] Bacaria J., Col J., M., Sánchez-Montijano E.: *The Labour Market in Spain: Problems, Challenges and Future Trends*, SIM Europe, POLICY BRIEF, 01 (2015)

### Kontaktné údaje

Ing. Tatiana Polonyankina.

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky  
nám. W. Churchilla 4, 130 67 Praha 3 – Žižkov

Tel: (+420) 774 871 148

email: [tatiana.polonyankina@vse.cz](mailto:tatiana.polonyankina@vse.cz)

# PRIESTOROVÁ DIFERENCIÁCIA A TEÓRIA HIER

## LOCATION MODELS IN THE GAME THEORY

*Allan Jose Sequeira Lopez*

### **Abstrakt**

Cieľom tohto projektu je analýza problematiky umiestnenia vzájomne si konkurujúcich firiem na nedokonalom trhu prostredníctvom skúmania lokačných hier Hotellinga. Na základe cenovej Hotellingovej hry analyzujeme problematiku umiestnenia firmy na nedokonalom trhu. V prípade trhu s homogénnym výrobkom zákazníci primárne preferujú dodávateľa podľa lokality, v ktorej žijú a len v prípade rozdielu v cene sú ochotní prejsť k inému.

*Kľúčové slová: lokácia, cena, Hotellingove hry, teória hier.*

### **Abstract**

The goal of the project is to analyse the location of competitive companies on the imperfect market using the Hotelling locations games. The issue of a company's location on the imperfect market is analyzed via Hotelling Pricing Game. In case of a market with homogenous product the customers tend to prefer the supplier according to their location, but, if prices diverge enough, they will be willing to switch from one supplier to the other.

*Keywords: Location, Pricing, Hotelling Games, Game Theory.*

## **1. MODEL PRIESTOROVEJ DIFERENCIÁCIE**

V lokačných modeloch predstavujú charakteristiky diferencovaných produktov body v priestore. Ak sa autá odlišujú len v spotrebe, priestor je jedno rozmerná línia. Ak je dôležité aj zrýchlenie, priestor je dvojrozmerná rovina. Tento prístup môžeme pochopiť, keď zoberieme do úvahy miesto, kde sa produkt predáva. Zákazníci môžu mať preferencie podľa toho, v akej oblasti žijú, budú teda uprednostňovať obchod, ktorý sa nachádza v ich blízkom okolí. Až v prípade, že sú ceny výrazne odlišné, zákazníci budú ochotní prejsť od jedného poskytovateľa k inému.

Harold Hotelling predstavil dva modely priestorovej diferenciacie – model voľby ceny a model voľby lokácie (HOTELLING, 1929). Na nasledujúcom obrázku je znázornené, čo majú oba modely spoločné. Dve firmy sú umiestnené v bode  $x_a$  a  $x_b$  pozdĺž línie, ktorá ide od 0 k 1 s konštantnou hustotou zákazníkov pozdĺž tejto línie. V cenovej hre podľa Hotellinga si firmy vyberajú ceny pre dané lokalizácie. V lokačnej hre podľa Hotellinga sú ceny fixné a firmy si vyberajú lokalizácie. V tomto príspevku sa venujeme lokačnému cenovému modelu.

### **1.1. CENOVÁ HRA**

V tejto hre sú hráčmi firmy A a B a kupujúci. Firmy sú umiestnené v  $x_a$  a  $x_b$ , kde  $x_a < x_b$ , a kupujúcim sa prideli index na základe umiestnenia  $x \in [0,1]$ . Priebeh hry je nasledujúci: firmy si simultánne vyberajú ceny  $p_a$  a  $p_b$  a každý kupujúci si vyberie firmu dodávateľa.

Dopyt je uniformne distribuovaný v intervale  $[0,1]$  s hustotou rovnou 1 (každý zákazník kupuje jednu jednotku). Produkčné náklady sú nulové. Každý zákazník vždy kúpi, takže jeho problémom je minimalizovať súčet ceny a lineárnych prepravných nákladov, ktoré sú  $\theta$  na prejdenú jednotku vzdialenosti, čo môžeme vyjadriť ako.

$$\pi_{kupujúci v x} = -\min\{\theta|x_a - x| + p_a, \theta|x_b - x| + p_b\}. \quad (1)$$

Na základe uvedeného uvažujeme o nasledujúcich možnostiach pre firmu A (analogicky platí pre firmu B):

$$\begin{cases} p_a(0) = 0 & \text{ak } p_a - p_b > \theta(x_b - x_a) \text{ firma B obsadí celý trh (a)} \\ p_a(1) = 1 & \text{ak } p_b - p_a > \theta(x_b - x_a) \text{ firma A obsadí celý trh (b)} \\ p_a\left(\frac{1}{2\theta}[(p_a - p_b) + \theta(x_a + x_b)]\right) & \text{v iných prípadoch, trh je rozdelený (c)} \end{cases} \quad (2)$$

Platby závisia od správania sa kupujúceho. Úžitok kupujúceho závisí od ceny, ktorú platí a od vzdialenosti, ktorú musí prejsť. Cenu teraz neberieme do úvahy. Firma A je najatraktívnejšia pre zákazníkov v  $x = 0$  („zákazník 0“) a najmenej atraktívna pre zákazníka v  $x = 1$  („zákazník 1“). Zákazník 1 kúpi od firmy A, pokiaľ bude platiť nasledujúca podmienka

$$\theta x_a + p_a < \theta x_b + p_b \quad (3)$$

Z čoho vyplýva, že

$$p_a - p_b < \theta(x_b - x_a) \quad (4)$$

Z čoho dostaneme zisk uvedený v (2a). Zákazník 1 kúpi od firmy B, ak

$$\theta(1 - x_a) + p_a > \theta(1 - x_b) + p_b \quad (5)$$

Čo implikuje

$$p_b - p_a < \theta(x_b - x_a) \quad (6)$$

Z čoho dostaneme zisk uvedený v (2b).

Je vysoko pravdepodobné, že nerovnosti v (4) a (6) sú obidve splnené, takže v tomto prípade zákazník 0 ide k firme A a zákazník 1 ide k firme B. V tomto prípade to reprezentuje zisk v (2c). Nasledujúcou úlohou je nájsť umiestnenie zákazníka  $x^*$ , ktorý je definovaný ako zákazník, ktorý je na hranici medzi dvoma trhmi, indiferentný od firmy A aj B.

V prvom rade si všimnime, že ak firma A pritiahne zákazníka  $x_b$ , môže tiež pritiahnuť všetkých  $x > x_b$ , pretože za  $x_b$  rastie vzdialenosť zákazníkov od oboch firiem rovnakou rýchlosťou. Takže vieme, že ak existuje indiferentný zákazník, bude sa nachádzať medzi  $x_a$  a  $x_b$ . Na základe tohto môžeme tvrdiť, že

$$\theta(x^* - x_a) + p_a = \theta(x_b - x^*) + p_b \quad (7)$$

Takže

$$p_b - p_a = \theta(2x^*x_a - x_b) \quad (8)$$

a

$$x^* = \frac{1}{2\theta} [(p_b - p_a) + \theta(x_a + x_b)] \quad (9)$$

Rovnica (9) platí len vtedy, ak skutočne existuje indiferentný zákazník. Ak takýto zákazník neexistuje, rovnica síce bude generovať výsledok pre  $x^*$ , ale toto číslo nebude mať význam.

Keďže firma A má všetkých svojich zákazníkov medzi 0 a  $x^*$ , rovnica (9) je dopytovou funkciou pre firmu A dovedy, až kým nestanoví svoju cenu tak vysoko nad firmou B, že stratí aj zákazníka 0. Dopyt, ktorému čelí firma B, je  $(1 - x^*)$ . Všimnime si, že ak  $p_b = p_a$ , potom z rovnice (9) platí  $x^* = \frac{x_a + x_b}{2}$ , nezávisle od  $\theta$ . Dopyt je lineárny pre ceny oboch firiem.

Po určení dopytových funkcií môžeme vypočítať Nashovu rovnováhu, keď si stanovíme ziskové funkcie pre každú firmu, deriváciou podľa ceny každej firmy a riešením dvoch podmienok prvého radu pre dve ceny. Ak existuje rovnovážny bod, v ktorom sú firmy ochotné vybrať také ceny, ktoré splnia nerovnosti (4) a (6), potom

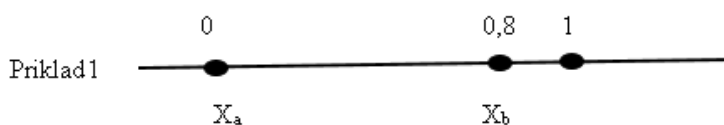
$$p_a = \frac{(2+x_a+x_b)\theta}{3}, p_b = \frac{(4-x_a+x_b)\theta}{3} \quad (10)$$

Z tejto rovnice môžeme vidieť, že firma A stanoví vyššiu cenu, ak väčšie  $x_a$  jej prinesie istejších zákazníkov alebo väčšie  $x_b$  zvýši počet konkurenčných zákazníkov. Najjednoduchší prípad je, keď  $x_a = 0$  a  $x_b = 1$ , kedy sa z rovnice dozvieme, že obidve firmy stanovujú cenu rovnú  $\theta$ .

### 1.1 Niekoľko príkladov s rôznymi variantmi

Vyskúšame  $x_a = 0$ ,  $x_b = 0,7$  a  $\theta = 0,5$ . Potom z rovnice (9) dostaneme  $p_a = (2+0+0,7)0,5/3=0,45$  a  $p_b = (4-0-0,7)0,5/3=0,55$ . Rovnica (10) hovorí, že

$$x^* = \frac{1}{2 * 0,5} [(0,55 - 0,45) + 0,5(0 + 0,7)] = 0,45$$

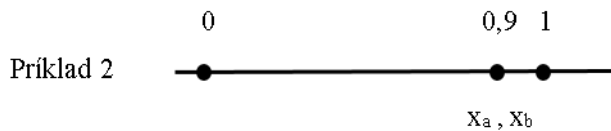


V príklade 1 existuje čisté strategické ekvilibium a rovnice generujú rozumné výsledky z vybraných parametrov.

### Príklad 2: Rovnaká lokalizácia ale iné ceny?

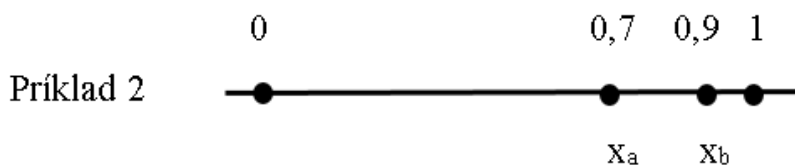
Vyskúšame  $x_a = 0,9$ ,  $x_b = 0,9$  a  $\theta = 0,3$ . Potom z rovnice (10) dostaneme  $p_a = (2+0,9+0,9)0,3/3 \approx 0,38$  a  $p_b = (4-0,9-0,9)0,3/3 \approx 0,22$ .

V tomto prípade rovnice generujú nezmyselné čísla, pretože ak sú obidve firmy sú umiestnené na 0,9 a  $p_a = 0,63$  a  $p_b = 0,37$ , firma B by obsadila celý trh. Tento výsledok nedáva zmysel, lebo rovnica (10) bola založená na predpoklade, že  $x_a < x_b$ , čo v tomto prípade neplatí.



**Príklad 3:  $x^* < x_a < x_b$ .**

Vyskúšame  $x_a = 0,7$ ,  $x_b = 0,9$  a  $\theta = 0,5$ . Potom z rovnice (10) dostaneme  $p_a = (2+0,7+0,9)0,5/3 = 0,6$  a  $p_b = (4-0,7-0,9)0,5/3 = 0,4$ . Čo sa týka rozdelenia trhu, rovnica (9) hovorí, že  $x^* = \frac{1}{2*0,5} [(0,4 - 0,6) + 0,5(0,7 + 0,9)] = 0,6$ .



Príklad 3 ukazuje na problém modelu. Ak sa trh rozdeľuje na  $x^* = 0,6$  ale  $x_a = 0,7$ ,  $x_b = 0,9$ , výsledok porušuje implicitný predpoklad, že hráči si rozdeľujú trh. Rovnica (9) je založená na predpoklade, že existuje nejaký indiferentný zákazník a ak je toto nesprávny predpoklad, ako v príklade 3, na základe rovnice (9) dostaneme hodnotu  $x^*$ , ale táto hodnota nebude mať zmysel. V skutočnosti zákazník na  $x = 0,6$  nie je skutočne indiferentný medzi firmou A a B. Môže si kúpiť od firmy A za celkové náklady  $0,6 + 0,1(0,5) = 0,65$  alebo od firmy B za  $0,4 + 0,3(0,5) = 0,55$ . V skutočnosti neexistuje žiadny zákazník, ktorý bude striktne uprednostňovať firmu A. Dokonca aj „domáci“ zákazník firmy A na hodnote  $x = 0,7$  by mal celkové náklady na nákup od firmy B  $0,4 + 0,5(0,9-0,7) = 0,5$  a teda by preferoval firmu B. Podobne aj zákazník na  $x = 0$  by mal celkové náklady na kúpu od firmy B vo výške 0,85, pri porovnaní s firmou A vy výške 0,95 by tiež preferoval firmu B.

Problémom v príklade 2 a 3 je, že firma s vyššou cenou by urobila lepšie, keby sa odchýlila svojou cenou tesne pod cenu konkurenta. Rovnica (10) bola vypočítaná s implicitným predpokladom, že lokálne maximum zisku je aj globálne maximum zisku, alebo inak: ak žiadna malá zmena nemôže zvýšiť zisk firmy, potom firma našla svoju optimálnu stratégiu. Niekedy veľká zmena zvýši hráčov zisk, aj keď malá zmena nie.



## ZÁVER

Na základe použitej literatúry sme pozornosť rôzne úskalia problematiky rozmiestnenia firiem z pohľadu viacerých vzájomne si konkurujúcich firiem na nedokonalom trhu. V prípade homogénneho produktu môžu mať zákazníci preferencie podľa lokality, v ktorej žijú. Len v prípade významného rozdielu v cene sú ochotní prejsť od jedného dodávateľa k inému. Uvedenú problematiku sme skúmali prostredníctvom analýzy Hotellingových hier: cenovej hry.

### Použitá literatúra

**EDGEWORTH, F. 1897.** *La Teoria Pura del Monopolio*. London : Macmillan, 1897.

**GOGA, M. (2013).** 2013. *Teória hier*. Bratislava : Iura Edition, 2013.

**HOTELLING, H. 1929.** *Stability in Competition*. s.l. : The Economic Journal, 1929.

**OSBORNE, M. P. 1986.** The Nature of Equilibrium in a Location Model. *International Economic Review*. 1986.

**PEKÁR J., B. I. 2012.** Modelovanie rozmiestňovania recyklačných centier. *Modelovanie rozmiestňovania recyklačných centier*. 2012.

**SHAKED, A. . 1982.** Existence and Computation of Mixed Strategy Nash Equilibrium for 3-Firms Location Problem. *The Journal of Industrial Economics*. 3 month, 1982, Zv. 31, 10.2307/2098006.

### Kontaktné údaje

Ing. Allan Jose Sequeira Lopez.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 824

email: sequeira\_36@hotmail.com

# VÝBĚROVÝ ROZPTYL NAD INTERVALOVÝMI DATY: SROVNÁNÍ OPTIMALIZAČNÍCH ALGORITMŮ V MATLABU

## SAMPLE VARIANCE OVER INTERVAL DATA: COMPARISON OF OPTIMIZATION ALGORITHMS IN MATLAB

*Ondřej Sokol*

### **Abstrakt**

V této práci jsou srovnávány tři univerzální algoritmy nelineární nekonvexní optimalizace dostupné v programu MATLAB se specializovaným algoritmem k výpočtu horní meze výběrového rozptylu nad intervalovými daty. Cílem je srovnat, jak přesné a jak časově náročné jsou tyto algoritmy, které obecně negarantují nalezení globálního optima. Srovnání je provedeno na třech různě generovaných typech dat a výsledné statistiky jsou vzájemně porovnány.

***Klíčová slova:** výběrový rozptyl, intervalová data, nelineární optimalizační algoritmy*

### **Abstract**

In this paper, we compare the universal algorithms for nonlinear nonconvex optimization (available in MATLAB) with specialized algorithm for computing upper bound of sample variance over interval data. The goal is to explore how accurate and how time consuming are chosen algorithms which do not guarantee global optimum. We focus on three different types of data and compare the resulting computational statistics.

***Keywords:** sample variance, interval, constrained nonlinear optimization algorithms*

## **1 INTRODUCTION**

Our goal is to explore how accurate and how time consuming are universal algorithms for nonconvex optimization (MATLAB implementations of interior point, active set and sequential quadratic programming methods) which do not guarantee global optimum in comparison to the specialized Ferson's algorithm in the case of computing upper bound of sample variance over interval data. We focus on three different types of data and compare the resulting computational statistics.

## **2 PROBLEM STATEMENT AND NOTATION**

Let us consider that a one-dimensional dataset  $x_1, \dots, x_n$  is unobservable; however, we know the collection of intervals

$$\mathbf{x}_1 = [x_1, \bar{x}_1], \dots, \mathbf{x}_n = [x_n, \bar{x}_n] \quad (1)$$

such that  $\underline{x}_i \leq x_i \leq \bar{x}_i$  for every  $i = 1, \dots, n$ .

The computation of the upper and lower bounds of sample variance, respectively, reduces to the optimization problems (2):

$$\begin{aligned}\underline{\hat{\sigma}} &= \min_{x \in R^n} \left\{ \frac{1}{n-1} \sum_{i=1}^n \left( x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 : \underline{x} \leq x \leq \bar{x} \right\}, \\ \bar{\hat{\sigma}} &= \max_{x \in R^n} \left\{ \frac{1}{n-1} \sum_{i=1}^n \left( x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 : \underline{x} \leq x \leq \bar{x} \right\}.\end{aligned}\tag{2}$$

Finding the lower bound (minimizing) is computable in polynomial time as it is a convex function over a convex set. Effective algorithms with the computational complexity of  $O(n \log n)$  can be found in [1]. As finding lower bound is possible in polynomial time (or better) no matter of data structure of dataset, it is not studied in this paper. We focus on the upper bound of sample variance.

The problem can be rewritten into standard quadratic optimization form:

$$\bar{\hat{\sigma}} = \max_{x \in R^n} \{x^T H x : \underline{x} \leq x \leq \bar{x}\}\tag{3}$$

where

$$H = \frac{n-1}{n} \begin{pmatrix} \frac{1}{n^2} - \frac{1}{n} & \frac{1}{n^2} & \dots & \frac{1}{n^2} \\ \frac{1}{n^2} & \frac{1}{n^2} - \frac{1}{n} & \dots & \frac{1}{n^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n^2} & \frac{1}{n^2} & \dots & \frac{1}{n^2} - \frac{1}{n} \end{pmatrix}\tag{4}$$

However, finding the upper bound (maximizing) is the NP-hard problem [2]. As it is the convex function over the convex set, at least one of the optimal solution is located at the edge of  $n$ -dimensional cube which is defined by the intervals.

In this paper we focus on the comparison of universal solvers with the specialized algorithm. We use MATLAB function `fmincon` for constrained nonlinear optimization with various parameters and specialized Ferson's algorithm [3]. The computational complexity properties of Ferson's algorithm was studied for example in [4–6].

### 3 DESCRIPTION OF THE ALGORITHMS

#### 3.1 Ferson's algorithm

This paper uses the Ferson's algorithm which was published in [3]. To fix notation, for an interval  $\mathbf{x} = [\underline{x}, \bar{x}]$  let  $x^C = \frac{1}{2}(\underline{x} + \bar{x})$  denotes its center and  $x^\Delta = \frac{1}{2}(\bar{x} - \underline{x})$  its radius. For  $\alpha > 0$ , let  $\alpha\mathbf{x}$  denote the  $\alpha$ -narrowed interval  $[x^C - \alpha x^\Delta, x^C + \alpha x^\Delta]$ . The algorithm works as follows:

1. We sort all  $2n$  endpoints of the narrowed intervals  $\frac{1}{n}\mathbf{x}_1, \dots, \frac{1}{n}\mathbf{x}_n$  into a sequence of ascending order  $x_{(1)}, \dots, x_{(2n)}$ . Then we divide the real line into  $2n + 1$  segments  $[x_{(k)}, x_{(k+1)}]$ , where we set  $x_{(0)} := -\infty$  and  $x_{(2n+1)} := \infty$ .
2. We compute  $\underline{E}$  and  $\bar{E}$  and select all intervals  $[x_{(k)}, x_{(k+1)}]$ , that intersect with  $[\underline{E}, \bar{E}]$ .
3. For each remaining interval  $[x_{(k)}, x_{(k+1)}]$  and for each  $i = 1, \dots, n$  we select the value of  $x_i$  following this rule:
  - if  $x_{(k+1)} < x_i^C - \frac{1}{n}x_i^\Delta$ , then  $x_i = \bar{x}_i$ ,
  - if  $x_{(k)} > x_i^C + \frac{1}{n}x_i^\Delta$ , then  $x_i = \underline{x}_i$ ,
  - otherwise we consider both possible values for  $x_i$ .
4. We get one or several sequences of  $x_i$ . If the average of  $x_1, \dots, x_n$  is in the interval of  $[x_{(k)}, x_{(k+1)}]$ , then we compute its sample variance using formula 
$$\hat{\sigma} = \frac{1}{n-1} \sum_{i=1}^n \left( x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2.$$
5. Return the largest value of the computed sample variances as  $\bar{\sigma}$ .

The exponential part of the computation complexity depends on the maximal number  $k$  of narrowed intervals having a common point. Then the computational complexity of the algorithm is  $O(n 2^k)$ . The algorithm gives the exact solution in the contrary to the following methods.

#### 3.2 Interior point method

Interior points method became popular in 1980s when Karmarkar [7] proposed his algorithm for linear programming. In the following years, there were many extensions of the algorithm to cover more general cases of problems – convex quadratic programming and nonconvex problems.

The main idea of this iterative algorithm is that the step from the feasible point is made towards the most effective direction. The length of the step must be chosen such that the resulting point is feasible. Steps are repeated until local optimum is found. For more details see for example [8] and [9].

In this paper, we use the MATLAB implementation (see [10]) of interior point algorithm to solve the problem. The algorithm in general does not guarantee finding the global optimum. Therefore, we run the algorithm multiple times from different starting points.

### 3.3 SQP and Active set methods

From the Matlab website [10] the main idea following: The Active set algorithm as well as the sequential quadratic programming algorithm transform the problem into an easier subproblems that can then be solved and used as the basis of an iterative process. SQP algorithm takes every iterative step in the region constrained by bounds. Furthermore, finite difference steps also respect bounds. Bounds are not strict; a step can be exactly on a boundary. This strict feasibility can be beneficial when your objective function or nonlinear constraint functions are undefined or are complex outside the region constrained by bounds.

The SQP algorithm combines the objective and constraint functions into a merit function. The algorithm attempts to minimize the merit function subject to relaxed constraints. This modified problem can lead to a feasible solution. However, this approach has more variables than the original problem, so the problem size increases [10].

For more information about these methods see [11] and [12]. In this paper, we use MATLAB implementations of these algorithms.

## 4 SIMULATION STUDY

The interior point, SQP and active set algorithms were run 10 times from different starting points with maximal number of iterations set to 5000 – this is not reached in our simulations. Tolerances are set to default value ( $10^{-16}$ ). The time includes all 10 instances with different starting points. Ferson’s algorithm is run only once as it is guaranteed to find the optimal solution. We ran every experiment 40 times.

Our goal is to compare the results and time of above mentioned method. It is expected that universal algorithms should be faster in problematic (e. g. higher number of intersecting narrowed intervals) instances; however, we need to explore how much differs the found solutions by theses algorithms which may be stuck in local extremes.

We focus on three types of datasets:

1. Large number of intervals and small number of narrowed interval intersection,
2. Medium number of intervals and medium number of narrowed interval intersection,
3. Small number of intervals and large number of narrowed interval intersection.

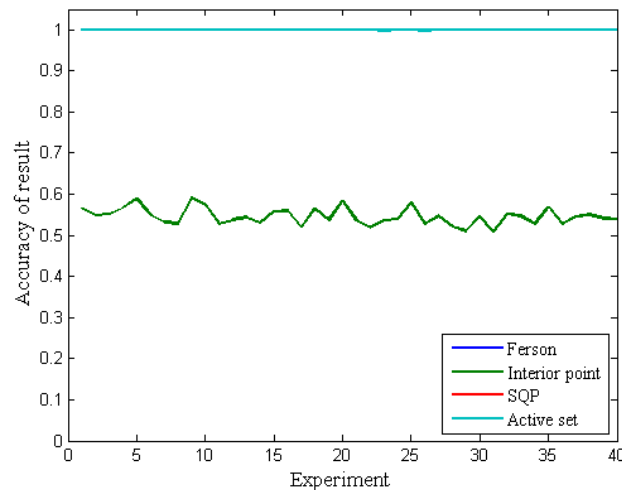
In particular, we used following parameters for generating data:

1.  $x^C \sim N(0,1), x^\Delta \sim Exp(1), n = 300$ , then  $\bar{k} = 4,72$ ,
2.  $x^C \sim N(0,1), x^\Delta \sim Exp(10), n = 100$ , then  $\bar{k} = 13,23$ ,
3.  $x^C \sim N(0,1), x^\Delta \sim Exp(100), n = 20$ , then  $\bar{k} = 18,12$ .

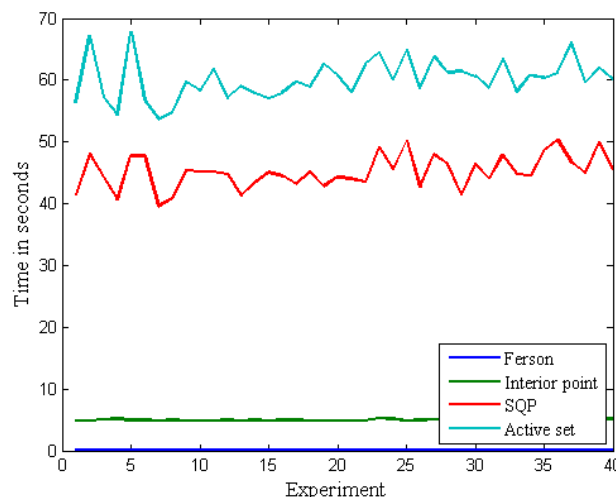
#### 4.1 Parameter set 1

The first simulation experiment was done with data generated under parameter set 1. Those data represent “common” situation as they are reasonable large and have not many intersections of narrowed intervals. This type of dataset suits the Ferson’s algorithm.

SQP and Active set algorithm often found the global optimum. At the average their solution was 0,9998 of the optimal solution with the worst case of 0.9983. However, they were also very time consuming in comparison to Ferson’s algorithm – it took almost 100 times longer to find solution from one starting point. The inferior point method was unable to find reasonable solution in this case. The results are shown in the Figure 1 and 2.



**Figure 1:** Found solution in comparison to global optimum, data generated under parameter set 1.



**Figure 2:** Time to find solution, data generated under parameter set 1.

#### 4.2 Parameter set 2

In the datasets generated with second parameter set is more intersection of narrowed intervals. As the result, the Ferson’s algorithm needs more time to find global optimum. The time varies considerably as with each intersection of narrowed intervals the number of iterations double.

In many cases, it is the most time consuming algorithm. The results of SQP and Active set algorithm are again very similar. They are not as close to global optimum as in previous case. In average, they find solution that is 0.996 of global optimum with the worst case of 0.983. SQP algorithm is the least time consuming of those three. The inferior point method was again unable to find reasonable solution in this case. The results are shown in the Figure 3 and 4.

### 4.3 Parameter set 3

In the last experiments, we have generated data that are inconvenient for the Ferson’s algorithm as there is many intersections of narrowed intervals in a small number of intervals. The consequence of such dataset is that Ferson’s algorithm needs almost 1000 times more time to find solution than other algorithms. However, SQP and Active set method are in this case not able to find global optimum or any close value to it with the average of 0.97 with the worst case of 0.87 in comparison to global optimum. Surprisingly the interior point method works better than SQP and Active set in this case with 0.993 average and worst case of 0.98. The results are shown in the Figure 5 and 6.

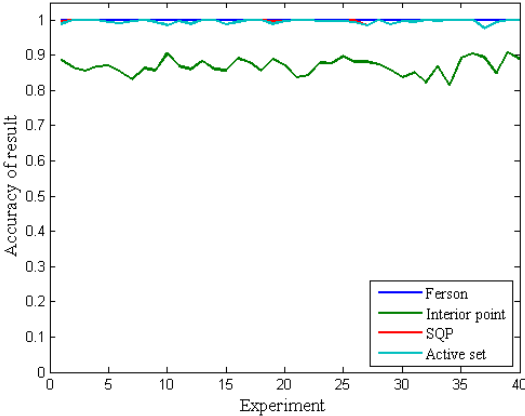


Figure 3: Found solution in comparison to the global optimum, data generated under parameter set 2.

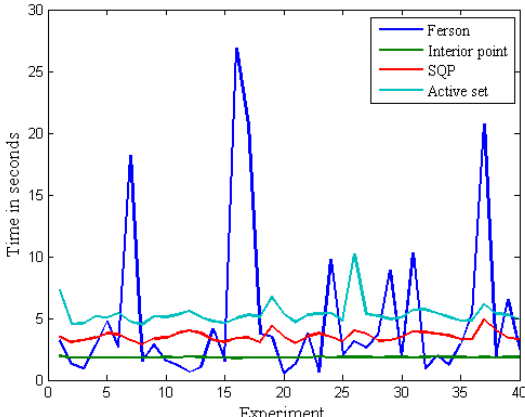
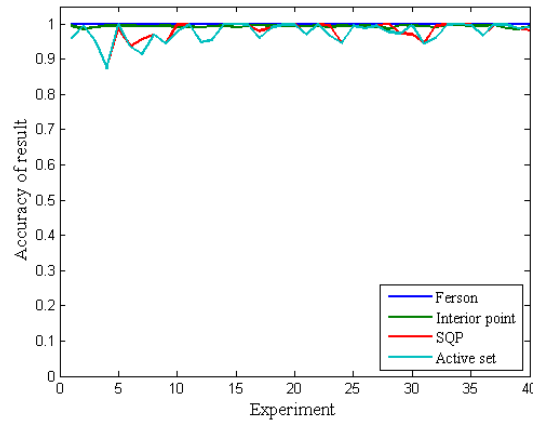
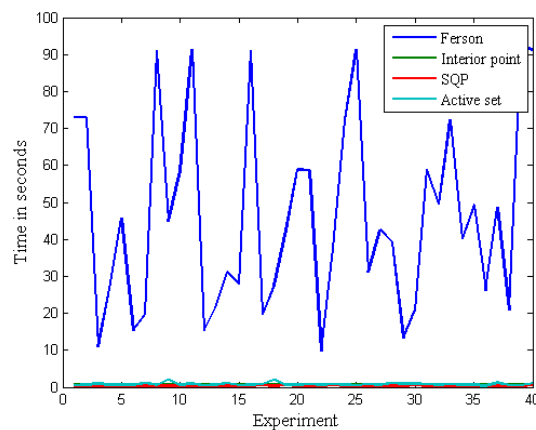


Figure 4: Time to find solution, data generated under parameter set 2.



**Figure 5:** Found solution in comparison to global optimum, data generated under parameter set 3.



**Figure 6:** Time to find solution, data generated under parameter set 3.

## 5 CONCLUSION

In this paper, we studied the accuracy and time computational complexity of four different methods finding the upper bound of the sample variance. As we expect, the only way to find the tight upper bound of sample variance is to use Ferson’s algorithm which was specially written for this purpose. Universal methods for nonlinear constrained optimization were surprisingly accurate but in most cases they did not found the global optimum but “good” local extremes. The analysis of accuracy and time consumption implies that it may be appropriate to use those methods if the tight bound is not needed and relatively close value is enough and the dataset is very inconvenient for the Ferson’s algorithm – if there are many intersections of narrowed intervals in relatively small number of intervals.

**Acknowledgements:** This work was supported by IGA F4/57/2017, University of Economics, Prague.



## References

- [1] XIANG, Gang, Martine CEBERIO and Vladik KREINOVICH. Computing Population Variance and Entropy under Interval Uncertainty: Linear-time Algorithms. *Reliable computing*. 2007, **13**(6), 467–488.
- [2] FERSON, Scott, Lev GINZBURG, Vladik KREINOVICH, Luc LONGPRÉ and Monica AVILES. Computing Variance for Interval Data is NP-hard. *ACM SIGACT News*. 2002, **33**(2), 108–118.
- [3] FERSON, Scott, Lev GINZBURG, Vladik KREINOVICH, Luc LONGPRÉ and Monica AVILES. Exact Bounds on Finite Populations of Interval Data. *Reliable Computing* [online]. 2005, **11**(3), 207–233. ISSN 1573-1340. Dostupné z: doi:10.1007/s11155-005-3616-1
- [4] SOKOL, Ondřej. Výpočet horní meze výběrového rozptylu nad intervalovými daty. In: *Nové trendy v ekonometrii a operačním výzkumu: Mezinárodní vědecký seminář nové trendy v ekonometrii a operačním výzkumu*. Bratislava: Ekonóm, 2015. ISBN 978-80-225-4181-7.
- [5] SOKOL, Ondřej and Miroslav RADA. How to Prove Polynomiality of Computation of Upper Bound Considering Random Intervals? In: *Mathematical Methods in Economics (MME2016)*. Liberec: TU Liberec, 2016, s. 779–785. ISBN 978-80-7494-296-9.
- [6] ČERNÝ, Michal and Ondřej SOKOL. Interval data and sample variance: A study of an efficiently computable case. In: *Mathematical Methods in Economics 2015 (MME)*. 2015.
- [7] KARMARKAR, N. A new polynomial-time algorithm for linear programming. In: [online]. B.m.: ACM Press, 1984, s. 302–311 [vid. 2017-08-30]. ISBN 978-0-89791-133-7. Dostupné z: doi:10.1145/800057.808695
- [8] FREUND, Robert M. and Shinji MIZUNO. Interior Point Methods: Current Status and Future Directions. In: Hans FRENK, Kees ROOS, Tamás TERLAKY a Shuzhong ZHANG, ed. *High Performance Optimization* [online]. Boston, MA: Springer US, 2000 [vid. 2017-05-18], s. 441–466. ISBN 978-1-4419-4819-9. Dostupné z: doi:10.1007/978-1-4757-3216-0\_18
- [9] POTRA, Florian A. and Stephen J. WRIGHT. Interior-point methods. *Journal of Computational and Applied Mathematics* [online]. 2000, **124**(1–2), 281–302. ISSN 03770427. Dostupné z: doi:10.1016/S0377-0427(00)00433-7
- [10] Constrained Nonlinear Optimization Algorithms. *MathWorks* [online]. [vid. 2017-08-30]. Dostupné z: <https://www.mathworks.com/help/optim/ug/constrained-nonlinear-optimization-algorithms.html>
- [11] BOGGS, Paul T. and Jon W. TOLLE. Sequential Quadratic Programming. *Acta Numerica* [online]. 1995, **4**, 1. ISSN 0962-4929, 1474-0508. Dostupné z: doi:10.1017/S0962492900002518
- [12] NOCEDAL, Jorge a Stephen J. WRIGHT. *Numerical optimization*. 2nd ed. New York: Springer, 2006. Springer series in operations research. ISBN 978-0-387-30303-1.

## Author's address

Ing. Ondřej Sokol

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky

náměstí Winstona Churchilla 4, 130 67 Praha 3

Czech Republic

Email: [ondrej.sokol@vse.cz](mailto:ondrej.sokol@vse.cz)

# PAVUČINOVÝ MODEL SO ZÁSOBAMI

## COBWEB MODEL WITH STOCKS

*Szomolányi Karol*

### Abstrakt

Trh tovaru, ktorý nepodlieha skaze a producenti ho môžu neobmedzene skladovať nie je stabilný. Vyplýva to z dynamickej analýzy pavučinového modelu so zásobami a bezprostredne realizovanými očakávaniami realizovanej v tejto práci.

*Kľúčové slová:* pavučinový model, zásoby, stabilita

### Abstract

The market of good, that is not perishable it may be stored by producers unlimitedly, is unstable. It follows from the dynamic analysis of the cobweb model augmented by stocks and immediate expectations realised in the paper.

*Keywords:* cobweb model, stocks, stability

## 1 ÚVOD

Ekonomické systémy sú dynamické. Túto skutočnosť by mali reflektovať ekonomické modely. Charakteristikou dynamického systému je, že ho možno opísať funkcionálnou rovnicou času (Gangolfo, 2010).

V tomto príspevku preveríme stabilitu pavučinového modelu charakterizujúceho cenu tovaru, ktorý nepodlieha skaze a producenti ho môžu neobmedzene skladovať.

## 2 MODEL

Uvažujme o pavučinovom modeli v tvare:

$$\begin{aligned}d_t &= a + bp_t \\s_t &= a_1 + b_1 p_t^e\end{aligned}\tag{1}$$

kde symbolmi  $d$  a  $s$  označujeme množstevný dopyt a ponuku a symbolmi  $p$  a  $p^e$  označujeme cenu a očakávanú cenu daného tovaru. Predpokladáme, že proces výroby tovaru vyžaduje nejaký čas a preto, producenti v čase iniciovania výroby odhadnú cenu tovaru v období  $t$ . Zvyšné symboly ( $a$ ,  $b$  ...) sú štrukturálne parametre lineárneho modelu. Predpokladáme, že tovar nepodlieha skaze a preto si producenti vytvárajú zásoby. Tvorba zásob podlieha špekulačnému mechanizmu v tvare:

$$z_t = \alpha(p_{t+1}^e - p_t); \alpha > 0\tag{2}$$

Ak producenti očakávajú pokles ceny v budúcnosti, znižujú zásoby, naopak ak predpovedajú rast ceny, zvyšujú zásoby. Podmienka rovnováhy na trhu je v tvare:

$$d_t + z_t = s_t + z_{t-1}\tag{3}$$

Predpokladáme, že očakávania sa realizujú bezprostredne a platí:

$$p_t^e = p_t \quad (4)$$

Kombináciou (1), (2), (3) a (4) možno model zapísať diferencnou rovnicou druhého rádu:

$$p_t - \frac{b_1 - b + 2\alpha}{\alpha} p_{t-1} + p_{t-2} = \frac{a_1 - a}{\alpha} \quad (5)$$

V ďalšom texte analyzujeme riešenie rovnice (5). Využívame postup Gandolfa (2010)

### 3 PARTIKULÁRNE RIEŠENIE A PODMIENKY STABILITY

Metódou neurčených koeficientov konštatujeme, že partikulárne riešenie rovnice (5) je:

$$\begin{aligned} \bar{p} &= \frac{a - a_1}{b_1 - b}, \text{ ak } b_1 \neq b \\ \bar{p}(t) &= \frac{a_1 - a}{2\alpha} t^2, \text{ ak } b_1 = b \end{aligned} \quad (6)$$

Prípado  $b_1 = b$  analyzujeme v .

Podmienky stability modelu opísaného diferencnou rovnicou (5) sú splnené ak platí:

$$\begin{aligned} 2 - \frac{b_1 - b + 2\alpha}{\alpha} &> 0 \\ 2 + \frac{b_1 - b + 2\alpha}{\alpha} &> 0 \\ 1 - 1 &> 0 \end{aligned} \quad (7)$$

Tretia podmienka (7) sa nespĺňa a preto vo všeobecnosti podmienky stability nie sú splnené. Vo všeobecnosti však nemožno konštatovať ani explozívny priebeh, pretože tretia podmienka nie je ostro porušená (ľavá strana sa rovná pravej,  $0 = 0$ ).

### 4 VŠEOBECNÉ RIEŠENIE

Všeobecné riešenie rovnice (5) sa rovná všeobecné riešenie prislúchajúcej homogénnej rovnice plus partikulárne riešenie. Prislúchajúcu charakteristickú rovnicu možno zapísať v tvare:

$$\lambda^2 - \frac{b_1 - b + 2\alpha}{\alpha} \lambda + 1 = 0 \quad (8)$$

kde symbolom  $\lambda$  označujeme charakteristický koreň (vlastné číslo). Charakteristická rovnica (8) je kvadratická a teda môže mať dva reálne odlišné korene, jeden násobný koreň alebo komplexne združené korene. Závisí od hodnoty diskriminanta kvadratickej rovnice:

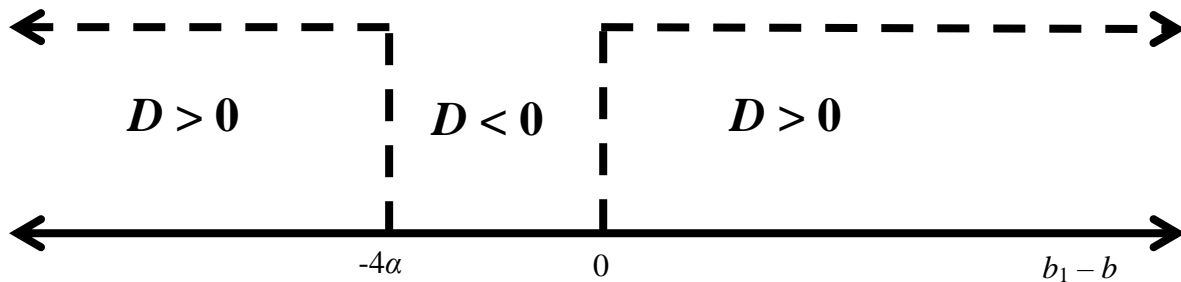
$$D = \left( -\frac{b_1 - b + 2\alpha}{\alpha} \right)^2 - 4 = \frac{(b_1 - b)(b_1 - b + 4\alpha)}{\alpha^2} \quad (9)$$

Charakteristické korene rovnice (8), ak

$$D = \frac{(b_1 - b)(b_1 - b + 4\alpha)}{\alpha^2} > 0 \Leftrightarrow (b_1 - b > 0) \vee (b_1 - b < -4\alpha) \quad (10)$$

Grafická interpretácia poznatku (9) je na obr. 1. Ak  $b_1 - b = -4\alpha$ , alebo  $b_1 - b = 0$ :  $D = 0$ . Poznamenajme, že v tom druhom prípade je partikulárne riešenie (6) kvadratickou funkciou času.

Obr. 1: Diskriminant charakteristickej rovnice (9)



Prameň: vlastná úvaha

Analyzujeme rôzne prípady.

#### 4.1 Dva odlišne reálne korene

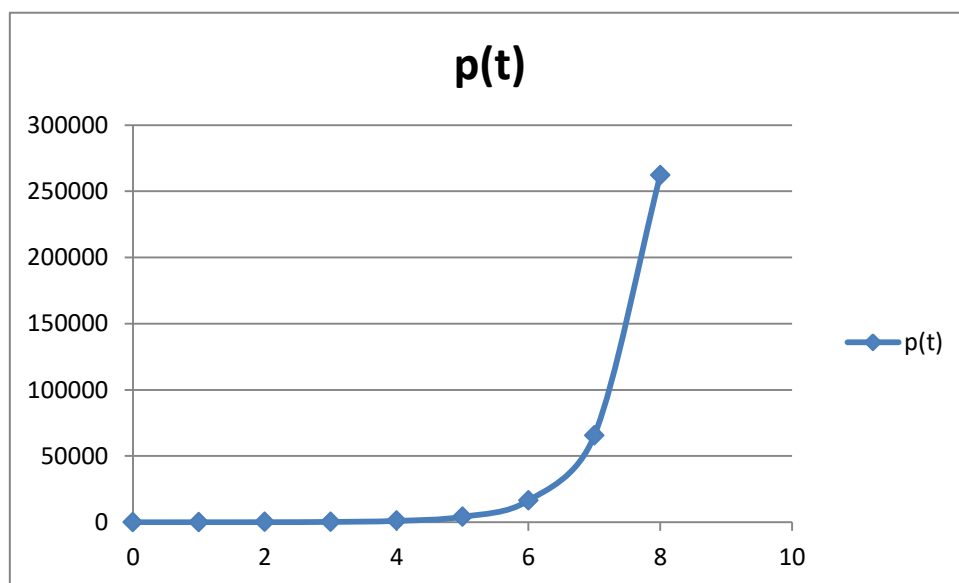
Charakteristické korene rovnice (8) sú dva odlišné reálne, ak  $b_1 - b < -4\alpha$ , alebo  $b_1 - b > 0$ . Ekonomicky zmysluplný je predpoklad,  $b_1 - b > 0$ , pretože predpovedáme kladný sklon ponuky  $b_1 > 0$  a záporný sklon dopytu  $b < 0$  (štandardný tovar). V prípade, ak  $b_1 - b < -4\alpha$ , druhá nerovnosť podmienok stability (7) je ostro porušená. V prípade  $b_1 - b > 0$  je zas ostro porušená druhá nerovnosť podmienok stability (7). Preto Priebeh ceny je v oboch prípadoch explozívny, nestabilný. Všeobecné riešenie diferenciálnej rovnice (5) je v tvare:

$$p(t) = A_1 (\lambda_1^*)^t + A_2 (\lambda_2^*)^t + \bar{p} \quad (11)$$

kde  $A_1$  a  $A_2$  sú ľubovoľné konštanty, hviezdičkou sú označené hodnoty charakteristických koreňov.

Na obr. 2 je graficky zobrazený priebeh ceny v čase v prípade  $b_1 = 4$ ,  $b = -4$ ,  $\alpha = 4$ ,  $a - a_1 = 8$ ,  $p_0 = 8/9$  a  $p_1 = 143/9$ .

Obr. 2: Dva odlišné reálne korene



Prameň: vlastné výpočty

## 4.2 Jeden reálny násobný

Charakteristický koreň rovnice (8) je jeden násobný, ak  $b_1 - b = -4\alpha$ , alebo  $b_1 - b = 0$ . V prípade, ak  $b_1 - b = -4\alpha$ , druhá nerovnosť podmienok stability (7) je porušená, ale nie ostro. V prípade  $b_1 - b = 0$  je neostro porušená druhá nerovnosť podmienok stability (7). V oboch prípadoch sa totiž charakteristický koreň v absolútnej hodnote rovná 1:

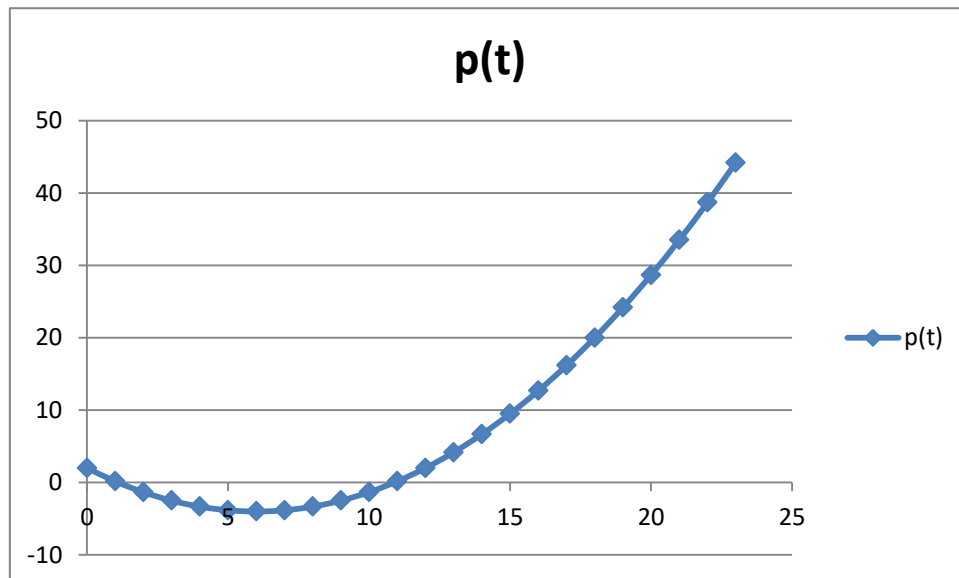
$$\begin{aligned}\lambda^* &= -1, \text{ ak } b_1 - b = -4\alpha \\ \lambda^* &= 1, \text{ ak } b_1 - b = 0\end{aligned}\quad (12)$$

Všeobecné riešenie diferenciálnej rovnice (5) je v tvare:

$$p(t) = (A_1 + A_2 t)(\lambda^*)^t + \bar{p} \quad (13)$$

Ak  $b_1 - b = -4\alpha$ , priebeh ceny je oscilačný ani konvergentný divergentný. Ak  $b_1 - b = 0$ , priebeh ceny bude explozívny, pretože partikulárne riešenie (6) je kvadratickou funkciou času. Druhý prípad s údajmi  $b_1 = 4$ ,  $b = 4$ ,  $\alpha = 4$ ,  $a - a_1 = -8$ ,  $p_0 = 2$  a  $p_1 = 1/6$  je zobrazený na obr. 3.

Obr. 3: Jeden reálny násobný koreň



Prameň: vlastné výpočty

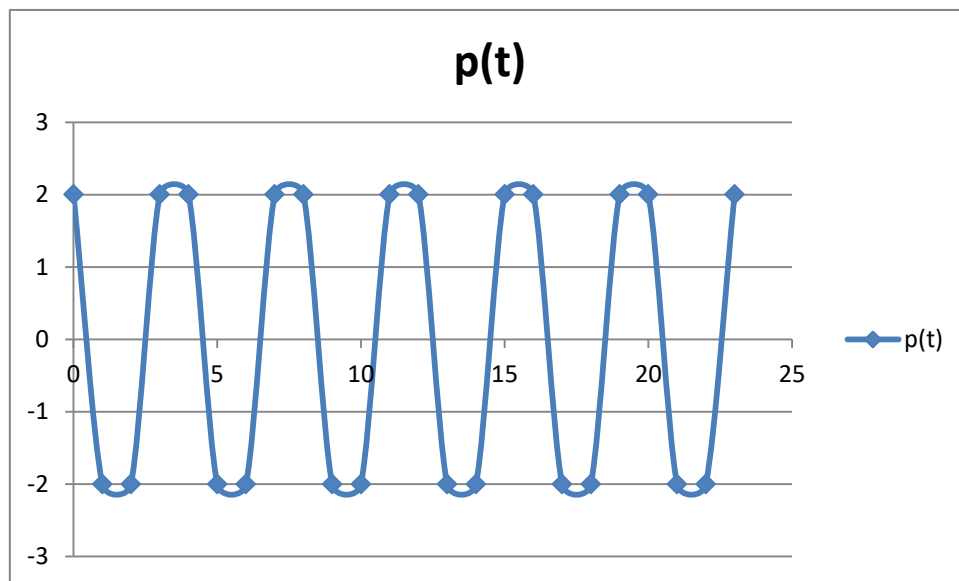
## 4.3 Komplexné združené korene

Charakteristické korene rovnice (8) sú komplexné združené, ak  $-4\alpha < b_1 - b < 0$ . Prvé dve nerovnosti podmienok stability sú splnené, avšak stále je neostro porušená tretia nerovnosť. Priebeh ceny bude oscilačný ani konvergentný ani divergentný. Všeobecné riešenie diferenciálnej rovnice (5) je v tvare:

$$\begin{aligned}p(t) &= A_1 \cos(\omega t) + A_2 \sin(\omega t) + \bar{p} \\ \cos \omega &= -\frac{1}{2}; \quad \sin \omega = \frac{1}{2}\sqrt{-D}\end{aligned}\quad (14)$$

Na obr. 4 je zobrazený priebeh ceny v čase pre prípad :  $b_1 = 4$ ,  $b = 12$ ,  $\alpha = 4$ ,  $a - a_1 = -8$ ,  $p_0 = 2$  a  $p_1 = -2$ .

Obr. 4: Komplexné združené korene



Prameň: vlastné výpočty

#### 4.4 Záver

Priebeh ceny tovaru nepodliehajúceho skaze, voľne zásobovateľného producentmi nie je explozívny iba v prípade, ak koreň charakteristickej rovnice (7) je násobný reálny a  $b_1 - b = -4\alpha$ , alebo korene charakteristickej rovnice (7) sú komplexné združené. Vo všeobecnosti však ani v týchto prípadoch systém nie je stabilný a priebeh ceny nie je ani konvergentný ani divergentný. V ostatných prípadoch, vrátane štandardného, v ktorom je sklon k dopytu záporný a sklon k ponuke kladný, je priebeh ceny explozívny.

#### Použitá literatúra

1. GANDOLFO, G. 2010. Economic Dynamics. Springer; 4th ed. ISBN: 978-3642038624

#### Kontaktné údaje

doc. Ing. Karol Szomolányi, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 822

email: [szomolan@euba.sk](mailto:szomolan@euba.sk)

# PŘEDZPRACOVÁNÍ DAT V OBLASTI RETAILU

## DATA PREPROCESSING IN RETAIL ENVIRONMENT

*Michal Švarc*

### **Abstrakt**

Předzpracování dat je časově nejnáročnější částí celého procesu datové analýzy. Tento článek se zabývá způsobem získávání informací o geografické poloze na základě nekompletních a chybových adres. V rámci tohoto článku byla vytvořena, popsána a implementována procedura, která využívá webovou službu společnosti Google k identifikaci chybových adresy a k získání dodatečných informací o poloze obchodů obchodního řetězce a jeho konkurentů. Tato procedura pomohla zlepšit úspěšnost geolokace přibližně o 20 procentních bodů ve srovnání s triviálním použitím webové služby na geokódování adresy.

***Klíčová slova:** předzpracování dat, datová integrace, otevřená data, Google APIs*

### **Abstract**

Data pre-processing is one of the most time-consuming part of data analysis process. This paper focuses on the way how to get geographical information from incomplete and burdened by errors addresses. In the scope of this paper was created, described and implemented a procedure which is using web service provided by Google company. This web service enabled the option to develop a procedure for getting geo-information from corrupted addresses of retailer's stores and also map retailer's competitors. Created procedure helped to increase precision of stores geolocation by 20 percentage points in comparison to simple usage of geocoding web service.

***Keywords:** data preprocessing, data integration, open data, Google APIs*

## 1 INTRODUCTION

More and more business decision makers in the retail environment nowadays rely on recommendations that are made by data analysts. Moreover, the proper usage of transactional data contributes to better understanding of the own business at any management level as well as to higher customer satisfaction.

The focus of this paper lies in data pre-processing in retail environment. The first part of this paper provides the basic summary of the data pre-processing tasks. This part is followed by practical showcase where is shown how to utilize Google Maps API for getting information about nearest competitor stores.

## 2 RELATED WORK

Already mentioned data pre-processing is one piece of the knowledge discovery process. (Han and Kamber 2011) It can be also described as a set of techniques which are used before an application of data mining methods. (García et al. 2016) Here should be mentioned the fact, that before starting any data pre-processing, it is needed to collect desirable data. The literature describes this step as data selection (Soibelman and Kim 2002). Especially for the corporate reality, it is typical that the business is controlled with the help of multiple core ERP (Enterprise Resource Planning) systems which are not usually intended to produce a data in the appropriate structure for further analysis.

Once the data are collected from back-office systems, the data pre-processing can be initiated. In this stage, the focus goes on the present data corruptions which would cause either incorrect outputs or failures of data mining procedures. Missing values, out of range values (outliers), impossible data combinations or other inconsistencies can be considered as corruptions. (García, Luengo, and Herrera 2016)

In terms of time consumption, the data pre-processing takes really large part of the whole knowledge discovery process. Munson's study shows that preparing of data usually consumes between 50% and 80% of practitioner's time devoted to knowledge discovery process. (Munson 2012) Due to this fact, there is an attempt to develop automatic or semi-automatic processes which would simplify the data preparation process for experienced user as well as provide a guidance for the non-expert user. As a good example of this attempt can be used the solution proposed by (Bilalli et al. 2016) which predicts the most appropriate transformations for selected data mining algorithm and used data set.

In general, the processing  $P$  modifies the raw real world data vectors  $X_{ik}$  into new data vectors  $Y_{ij}$

$$Y_{ij} = P(X_{ik}),$$

where the new data vectors  $Y_{ij}$  contains "valuable informations" from  $X_{ik}$  and fixes at least one data corruption problem from  $X_{ik}$ . In above mentioned equation the index  $i = 1, \dots, n$  denotes the index of record and the number of records in data set is  $n$ . Index  $k$  represents number of features or attributes before data pre-processing and  $j$  relates to count of attributes after data pre-processing (Famili et al. 1997)

Data pre-processing that was mentioned above can be classified into these groups which focus on particular steps before the data can be used for data mining. These groups are:

- cleaning
- integration



- transformation
- reduction

## **2.1 Data cleaning**

Real-world datasets require special treatments that focus on incompleteness, noise and inconsistency. The human mistakes are just one possible source of these all "errors". In general, the data are generated by processes not primarily intended for data analysis and do not incorporate routines which would fill in missing values, identify missing values and correct discrepancies in the data.

Strategies for handling of missing values involve can be split into two branches. One branch simply ignores tuples with missing attributes. The second branch tries to substitute missing values. This substitution can be done either manually or automatically. Automated missing values substitution in its simplest form just labels missing as "Unknown". More advanced methods trying to substitute missing values with global or local measure of central tendency (e.g., the mean or median for quantitative variables) or attempts to estimate the most probable value. However, the intention of filling missing values entails the risk data bias. (Han and Kamber 2011)

Noise in data is represented by random errors or variances in measured variables. For example, outliers can be considered as noise in data. Commonly used ways of noise fixing are called smoothing methods which are based on bindings, regression and clustering or outlier analysis. (Han and Kamber 2011)

## **2.2 Data integration**

Common practice shows that there are always many data sources which provide relevant data for required analysis and need to be merged before the analysis can start. Nowadays, analysis of own data is just one piece of mosaic of reality. More and more data analysts calling for utilization of data from external sources. Combination of multiple data sources brings issues like mapping entities identifiers, detecting and resolving data value conflicts, handling duplicities. (Han and Kamber 2011)

## **2.3 Data transformation**

The main aim of the data transformation is to help data mining processes be more efficient in patterns searching. One of the data transformation strategies was already mentioned in data cleaning section and relates to smoothing methods. Another set of methods (normalization, discretization) are used for transformation of scales. Last but not least data transformation is represented by aggregations. For example, this helps to analyse retailer's categories performance, where a performance of products is aggregated into whole categories. Finally, there is an option to extend tuples with new attributes which are derived from their original characteristics. (Han and Kamber 2011)

## **2.4 Data reduction**

The motivation for data reduction may have following reasons. Data contains too many rows/tuples that data mining procedures would fail or their runtime would be too long. Second reason for data reduction is huge number of available attributes. Naturally, combination of both above mentioned reason can be a reason for data reduction.

First possible strategy for reducing data volume is a removal of unimportant attributes. If this approach cannot be used then number of rows can be easily reduced with help of sampling. But,

there is a risk that simple random sampling may have negative impact on results if there is a presence of skew. To minimize the impact, it may be useful to utilize stratified sampling. An alternative to sampling data can be used clustering. In this case the analysis is applied for each data cluster independently and the analysis conclusions may or may not be aggregated to the whole dataset. Data with high dimensionality can be reduced with the help of Feature Extraction methods like:

- Principal component analysis
- Factor analysis
- Independent component analysis
- etc...

An alternative to above mentioned Feature Extraction methods are heuristic methods which apply step-by-step approach for selection or otherwise eliminating used attributes. (Maimon and Rokach 2010)

### **3 PRACTICAL DEMONSTRATION**

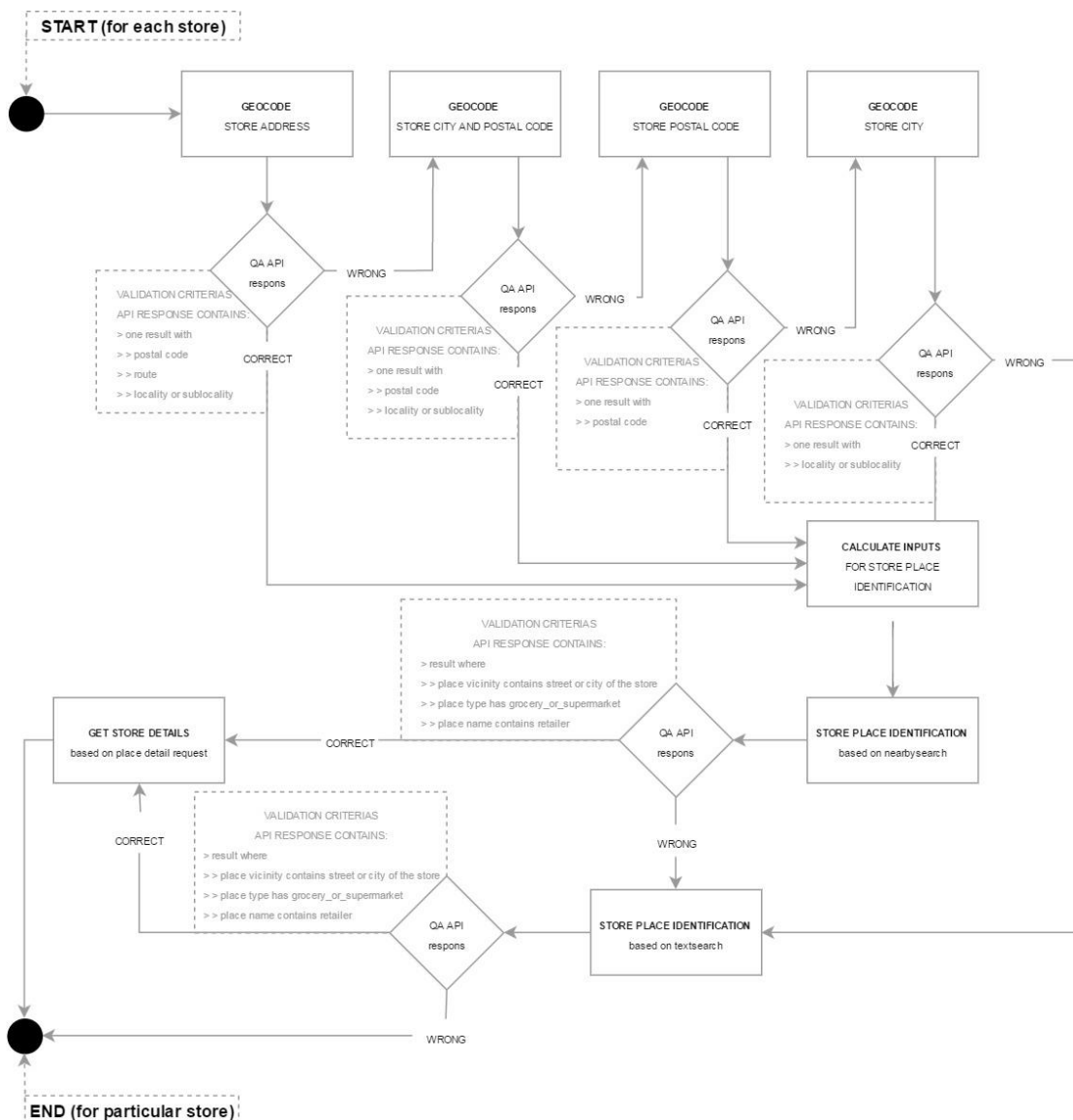
This section, practical demonstration, contains showcase of data integration part as a part of data preprocessing that uses external data source for enrichment of the analysts data warehouse. As already outlined above, the main aim is to obtain data about stores geoattributes like geographic coordinates and territorial administrative units based on incomplete store address. More over get the information about competitors stores in the vicinity of each retailer store.

Proposed solution is a procedure which is built up around the Google Geocoding API. ("Geocoding Service Google Maps JavaScript API" 2017) This service is mainly used for receiving geo coordinates based on the store address and listing supermarkets and stores around the given place. The main issue in this approach is represented by the store's addresses corruptions. These corruptions are caused by two factors. The first factor is represented by the human errors in typing addresses and the second set of errors occurs for all accented letters. Following line represent one of the worst tuple for store address: ?elezni???? 8/1303 Hav??ov 73601\* that is a representation of Železničářů 8/1303, Havířov, 736 01.

In terms of the automatization, following ETL process was created and deployed. The first stage of the ETL process checks the store dimension for actively trading stores in last two years. Outputs of this check becomes a inputs for a next stage where the core procedure starts and takes the responsibility for two major tasks. The first task related to getting geo attributes for actively trading stores without this information and the second job of the core procedure relates updating the list of competing stores. Final stage of the process sends information about newly obtained information to data warehouse owner for approval. Once the approval is done, the data are productionalized in the data warehouse.

At this place, there is a need to describe the first task of the core procedure in more details because it differs from the common way of API usage where the requestor gets the answer if the request is formulated correctly. This difference is caused by the store addresses corruptions and requires gradual approach.

Following diagram shows the gradual approach in details.

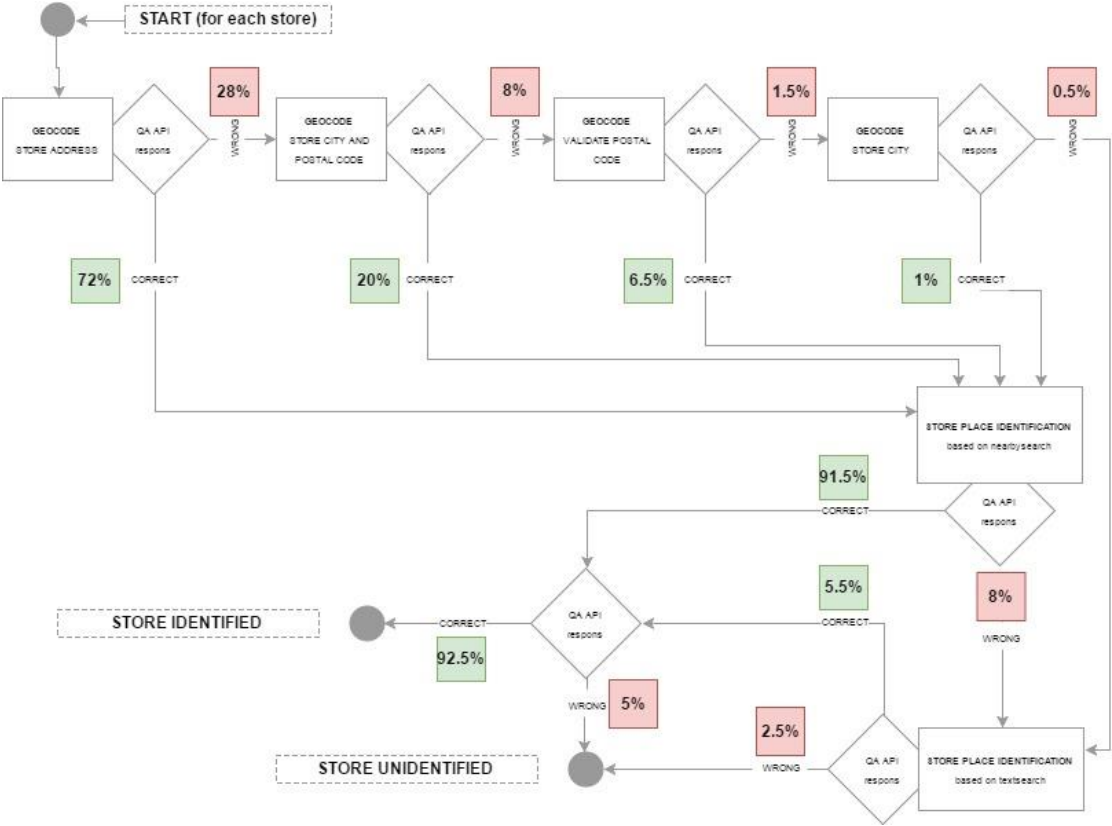


First of all, the core procedure tries to estimate the districts based on given store address where the store can be located. The district estimator uses Google maps geocoding API and can be over with two statuses (found district or district not found). If the district was found the procedure extracts the notional centre of the district where is the store located and calculates the radius around the notional centre of the district. These two attributes become inputs for the second step of the procedure. In this step, the core procedure push as inputs these two attributes with a retailer name and keyword `grocery_or_supermarket` against the Google's place nearbysearch API. If the google place nearbysearch response does not contain the retailer's store or if the district was not found the procedures push the request with the retailer name, store city and store street against the Google Place Textsearch API. The core procedure ends for a particular store with failure if the Google Place Textsearch API is not able to identify the store place. Once the Google Place Nearbysearch or Textsearch API returns a place id, the core procedure goes to the final where the Google Place Detail API is used for getting opening hours, ratings and parts of store address.

The core procedure was used for identification of 1093 actively trading Tesco stores in Czechia, Hungary, Poland and Slovakia. The following table summarizes the performance of core procedure and compares its results with the open-and-shut approach where the store geo attributes are obtained just with the help of simple store address geocoding.

	Core procedure identification		Open-and-shut identification	
	Count	Percentage	Count	Percentage
Total number of searched stores	1093	100.0%	1093	100.0%
Number of correctly identified stores	1012	92.6%	788	72.1%
Number of wrongly identified stores	26	2.38%	305	27.9%
Number of unidentified stores	55	5.03%		

Moreover, the next diagram shows the usage of individual procedure steps and their success rates.



#### 4 CONCLUSIONS AND SPACE FOR FURTHER WORK

Beside the brief introduction in the first part of this paper, the main added value of the whole work lies in the practical part. The procedure for getting store geo-attribute provided some reduction of manual analysts' work. Moreover, the proposed procedure flow reduced total costs for getting required geo-attributes, because of the elimination of text-search queries. Nowadays, Google company launched the extension of their maps API which is called Places autocomplete that focus on providing a Google's place ID based on given keywords and locality. It would be interesting to see the difference in performance between this new application and Places autocomplete API.

## Literature

1. BILALLI, Besim, Alberto ABELLÓ, Tomas Aluja-BANET, and Robert WREMBEL. 2016. "Automated Data Pre-Processing via Meta-Learning (PDF Download Available)." *ResearchGate*.  
[https://www.researchgate.net/publication/307872762\\_Automated\\_Data\\_Pre-processing\\_via\\_Meta-learning](https://www.researchgate.net/publication/307872762_Automated_Data_Pre-processing_via_Meta-learning)
2. FAMILI, A., Wei-Min SHEN, Richard WEBER, and Evangelos SIMOUDIS. 1997. "Data Preprocessing and Intelligent Data Analysis." *Intelligent Data Analysis* 1 (1): 3–23. doi:[10.1016/S1088-467X\(98\)00007-9](https://doi.org/10.1016/S1088-467X(98)00007-9)
3. GARCÍA, Salvador, Julián LUENGO, and Francisco HERRERA. 2016. "Tutorial on Practical Tips of the Most Influential Data Preprocessing Algorithms in Data Mining." *Knowledge-Based Systems* 98 (April): 1–29. doi:[10.1016/j.knosys.2015.12.006](https://doi.org/10.1016/j.knosys.2015.12.006).
4. GARCÍA, Salvador, Sergio RAMÍREZ-GALLEGO, Julián LUENGO, José Manuel BENÍTEZ, and Francisco Herrera. 2016. "Big Data Preprocessing: Methods and Prospects." *Big Data Analytics* 1: 9. doi:[10.1186/s41044-016-0014-0](https://doi.org/10.1186/s41044-016-0014-0).
5. HAN, Jiawei, and Micheline KAMBER. 2011. *Data Mining: Concepts and Techniques*. 3rd ed. Burlington, MA: Elsevier.
6. MAIMON, Oded, and Lior ROKACH, eds. 2010. *Data Mining and Knowledge Discovery Handbook*. Boston, MA: Springer US.  
<http://link.springer.com/10.1007/978-0-387-09823-4>.
7. MUNSON, M. Arthur. 2012. "A Study on the Importance of and Time Spent on Different Modeling Steps." *SIGKDD Explor. Newsl.* 13 (2): 65–71. doi:[10.1145/2207243.2207253](https://doi.org/10.1145/2207243.2207253).
8. SOIBELMAN, Lucio, and Hyunjoo KIM. 2002. "Data Preparation Process for Construction Knowledge Generation Through Knowledge Discovery in Databases." *Journal of Computing in Civil Engineering* 16 (1): 39–48. doi:[10.1061/\(ASCE\)0887-3801\(2002\)16:1\(39\)](https://doi.org/10.1061/(ASCE)0887-3801(2002)16:1(39)).

## Dedikace:

VŠE IGA F4/93/2017

## Contact

Ing. Michal Švarc.

University of Economics, Prague, Faculty of Informatics and Statistics

Faculty of Informatics and Statistics

Phone: 00420 724 031 631

Email: xsvam00@vse.cz

# POROVNÁNÍ ODHADŮ STOCHASTICKÉ VOLATILITY: ANALÝZA S&P 500

## COMPARISON OF STOCHASTIC VOLATILITY ESTIMATES: ANALYSIS OF S&P 500

*Petra Tomanová*

### Abstrakt

Cílem této práce je porovnání různých odhadů volatility výnosů akcií S&P 500 za použití modelu stochastické volatility. První jsou parametry modelu odhadnuty pomocí Kalmanova filtru. Poté je volatilita odhadnuta pomocí metody Kalman smoother a Importance sampling. Odhady jsou porovnány na výnosech S&P 500 v letech 2005 až 2014.

*Klíčová slova:* Stochastická volatilita, S&P 500, Kalman filter smoother, Importance sampling

### Abstract

The aim of this paper is to compare different volatility estimates of returns of S&P 500 based on stochastic volatility model. First, the parameters are estimated using Kalman filter. Then the volatility is estimated using Kalman smoother method and Importance sampling method. The approaches are compared using S&P 500 returns of the period which spans from 2005 to 2014.

*Keywords:* Stochastic Volatility, S&P 500, Kalman filter smoother, Importance sampling

## 1 ÚVOD

Modely stochastické volatility mají své uplatnění v oblasti financí a to například v ohodnocování opcí. Jejich základní charakteristikou je fakt, že volatilita je modelována jako nepozorovaná latentní proměnná. Přehled modelů stochastické volatility lze například nalézt v [2]. V této práci je aplikován model stochastické volatility na výnosech akcií S&P 500. Cílem je srovnat výsledky za použití Kalman filter smoother a Importance sampling. Práce navazuje na článek [4].

## 2 MODELOVÁNÍ FINANČNÍCH CEN A VÝNOSŮ

### 2.1 Model stochastické volatility

Nechť  $P_t$  značí uzavírací cenu obchodního dne  $t$ . Pak denní výnos  $r_t$  lze spočítat jako

$$r_t = \log(P_t/P_{t-1}) = \Delta \log P_t = \Delta p_t.$$

$p_t$  lze chápat jako diskretizovanou realizaci spojitého procesu logaritmické ceny  $\log P_t$ ,

$$d \log P(t) = \mu dt + \sigma(t) dW(t),$$

kde  $\mu$  je střední hodnota výnosu,  $\sigma(t)$  je spojitý proces volatility a  $W(t)$  je standardizovaný Brownův pohyb. Nechť  $\log \sigma(t)^2$  je generován Ornstein-Uhlenbeckovým procesem

$$\log \sigma(t)^2 = \xi + H(t), \quad dH(t) = -\lambda H(t) dt + \sigma_\eta dB(t),$$

kde  $\xi$  je konstanta,  $0 < \lambda < 1$ ,  $\sigma_\eta$  je volatilita volatility (striktně pozitivní) a  $B(t)$  je standardizovaný Brownovský pohyb nezávislý na  $W(t)$  [1].

V souladu s [1] za použití Euler-Maruyama diskretizační metody lze pro denní výnosy  $y_t$  specifikovat model stochastické volatility následovně

$$y_t = \mu + \sigma_t \varepsilon_t, \quad \log \sigma_t^2 = \xi + H_t, \quad H_{t+1} = \phi H_t + \sigma_\eta \eta_t,$$

kde  $\phi = 1 - \lambda$ ,  $0 < \phi < 1$ . Tento nelineární model lze transformovat pomocí substituce  $x_t = \log(y_t - \mu)$  na

$$x_t = h_t + u_t, \quad h_{t+1} = \omega + \phi h_t + \sigma_\eta \eta_t,$$

kde  $u_t = \log \varepsilon_t^2$ ,  $\omega = (1 - \phi)\xi$  a  $h_t = H_t + \xi$ .

Stochastický model pro časové řady finančních výnosů  $y_t$  s  $\mu = 0$  lze tedy zapsat jako

$$y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \exp(h_t), \quad \varepsilon_t \sim \text{NID}(0, 1),$$

kde logaritmická volatilita  $h_t$  sleduje stacionární autoregresní proces

$$h_{t+1} = \omega + \phi h_t + \sigma_\eta \eta_t, \quad \eta_t \sim \text{NID}(0, 1)$$

pro  $t = 1, \dots, n$  s  $h_1 \sim N(\omega/1 - \phi, \sigma_\eta^2/1 - \phi^2)$ , se vzájemně a sériově nezkorelovanými  $\varepsilon_t$  a  $\eta_t$ .

## 2.2 Kalman filter smoother

Mějme lineární Gaussovský state space model definován podle [3], [1]

$$\begin{aligned} \alpha_{t+1} &= T_t \alpha_t + R_t \zeta_{1t}, & \zeta_{1t} &\sim \text{NID}(0, Q_t), \\ y_t &= Z_t \alpha_t + \zeta_{2t}, & \zeta_{2t} &\sim \text{NID}(0, H_t) \end{aligned}$$

s počáteční podmínkou  $\alpha_1 \sim N(a_1, P_1)$ .  $\zeta_{1t}$  a  $\zeta_{2s}$  jsou nezávislé pro všechny  $t, s$  a nezávislé od  $\alpha_1$ . Vektor  $\alpha_t$  je nepozorovaný a matice  $T_t, Z_t, R_t, Q_t, H_t$  jsou fixní v čase  $t$ .

Nepozorovaný vektor  $\alpha_t$  lze odhadnout z pozorování za použití Kalmanova filtru [1]:

$$\begin{aligned} v_t &= y_t - Z_t \alpha_t, \\ F_t &= Z_t P_t Z_t' + H_t, \\ K_t &= T_t P_t Z_t' F_t^{-1}, \\ a_{t+1} &= T_t a_t + K_t v_t, \\ P_{t+1} &= T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t' \end{aligned}$$

pro  $t = 1, \dots, n$  a počáteční hodnoty  $a_1$  a  $P_1$ . Platí  $Y_t = \{y_1, \dots, y_t\}$  a dále

$$a_{t+1} = \mathbb{E}(\alpha_{t+1} | Y_t), \quad P_{t+1} = \mathbb{V}\text{ar}(\alpha_{t+1} | Y_t).$$

Poté co jsou filtrované odhady spočítány, lze použít následující rekurzi (Kalman smoother) začínající posledním pozorováním a končící prvním [1]:

$$\begin{aligned} r_{t-1} &= F_t^{-1} v_t + L_t r_t, \\ N_{t-1} &= F_t^{-1} + L_t' N_t L_t, \\ \hat{\alpha}_t &= a_t + P_t r_{t-1}, \\ V_t &= P_t - P_t N_{t-1} P_t. \end{aligned}$$

Pro výpočet odhadu vyrovnaného modusu  $h_t$  lze použít následující rovnice vycházející z Newton-Raphson algoritmu

$$A = \text{diag} \frac{2 \times \exp(h)}{Y_n} \Big|_{h=g}, \quad z = g - \frac{1}{2}A + \iota, \text{ where } \iota = (1, 1, \dots, 1)^T.$$

Položením  $H = A$  a  $Y_n = z$  a za použití Kalman filter smoother lze pak získat odhady modusu.

### 2.3 Importance sampling

Cílem je odhadnout podmíněnou střední hodnotu  $\bar{x}_t = \mathbb{E}[x_t(\alpha_{1:t})|Y_t]$ . Pro dané  $\psi$ , necht'  $g(\alpha_{1:t}|Y_t)$  je gaussianská "importance" hustota:

$$\begin{aligned} \bar{x}_t &= \int x_t(\alpha_{1:t}) p(\alpha_{1:t}|Y_t) d\alpha_{1:t} \\ &= \int x_t(\alpha_{1:t}) \frac{p(\alpha_{1:t}|Y_t)}{g(\alpha_{1:t}|Y_t)} g(\alpha_{1:t}|Y_t) d\alpha_{1:t} \\ &= \mathbb{E}_g \left[ x_t(\alpha_{1:t}) \frac{p(\alpha_{1:t}|Y_t)}{g(\alpha_{1:t}|Y_t)} \right], \end{aligned}$$

kde  $\mathbb{E}_g$  značí střední hodnotu vzhledem k "importance" hustotě  $g(\alpha_{1:t}|Y_t)$ . Jelikož  $p(\alpha_{1:t}, Y_t) = p(Y_t)p(\alpha_{1:t}|Y_t)$ , pak

$$\bar{x}_t = \mathbb{E}_g \left[ x_t(\alpha_{1:t}) \frac{p(\alpha_{1:t}|Y_t)}{g(\alpha_{1:t}|Y_t)} \right] = p(Y_t)^{-1} \mathbb{E}_g [x_t(\alpha_{1:t}) \tilde{w}_t],$$

kde

$$\tilde{w}_t = \frac{p(\alpha_{1:t}, Y_t)}{g(\alpha_{1:t}|Y_t)}.$$

Položením  $x_t(\alpha_{1:t}) = 1$  dostaneme  $p(Y_t) = \mathbb{E}_g(\tilde{w}_t)$  a

$$\bar{x}_t = \frac{\mathbb{E}_g [x_t(\alpha_{1:t}) \tilde{w}_t]}{\mathbb{E}_g(\tilde{w}_t)}.$$

Monte Carlo odhad  $\bar{x}_t$  je pak dán rovnicí

$$\hat{x}_t = \sum_{i=1}^N x_t(\alpha_{1:t}^{(i)}) w_t^{(i)},$$

kde

$$\tilde{w}_t^{(i)} = \frac{p(\alpha_{1:t}^{(i)}, Y_t)}{g(\alpha_{1:t}^{(i)}|Y_t)}, \quad w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^N \tilde{w}_t^{(j)}}$$

a  $\alpha_{1:t}^{(i)}$  je nezávislý výběr z  $\alpha_{1:t}^{(i)} \sim g(\alpha_{1:t}|Y_t)$ ,  $i = 1, \dots, N$ . Řada  $w_t^{(1)}, \dots, w_t^{(N)}$  se skládá z normalizovaných importance sampling vah. [1]

## 3 ANALÝZA S&P 500

### 3.1 Data

V této práci jsou analyzovány ceny akcií S&P 500 od začátku roku 2005 do konce roku 2014. Upravené uzavírací ceny byly použity z článku [4] (původní zdroj: [www.finance.yahoo.com](http://www.finance.yahoo.com)). Pro účely analýzy byly z cen akcií spočteny procentuální logaritmičké výnosy. Poté byla časová řada transformována pomocí substituce za účelem linearizace modelu, neboli  $x_t = \log(y_t - \mu)^2$ .



### 3.2 Výsledky

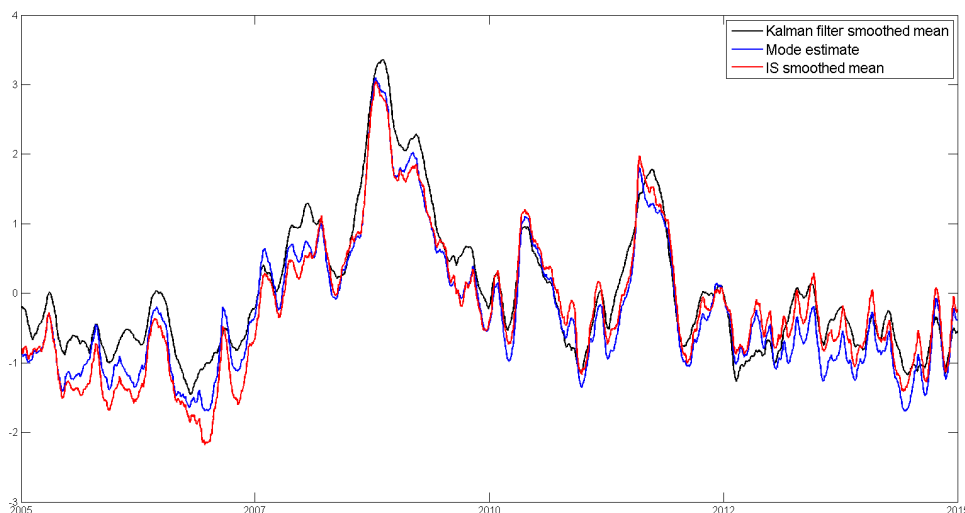
Pro odhady koeficientů se předpokládá, že  $u_t = \log \varepsilon_t^2 \sim \text{NID}(0, \sigma_u^2)$  a  $\eta_t \sim \text{NID}(0, \sigma_\eta^2)$ . Neznámé koeficienty  $\sigma_u^2$ ,  $\sigma_\eta^2$ ,  $\omega$  a  $\phi$  lze pak odhadnout metodou Quasi Maximum Likelihood, kde log-likelihood je získána z Kalmonového filtru. Omezení parametrů bylo stanoveno následovně:  $\sigma_u^2 > 0$ ,  $\sigma_\eta^2 > 0$ ,  $\omega > 0$  a  $0 < \phi < 1$ . Hodnoty odhadnutých koeficientů jsou následující:

$$\sigma_u^2 = 5.7399, \quad \sigma_\eta^2 = 0.0180, \quad \omega = 1.0011e - 07, \quad \phi = 0.9977,$$

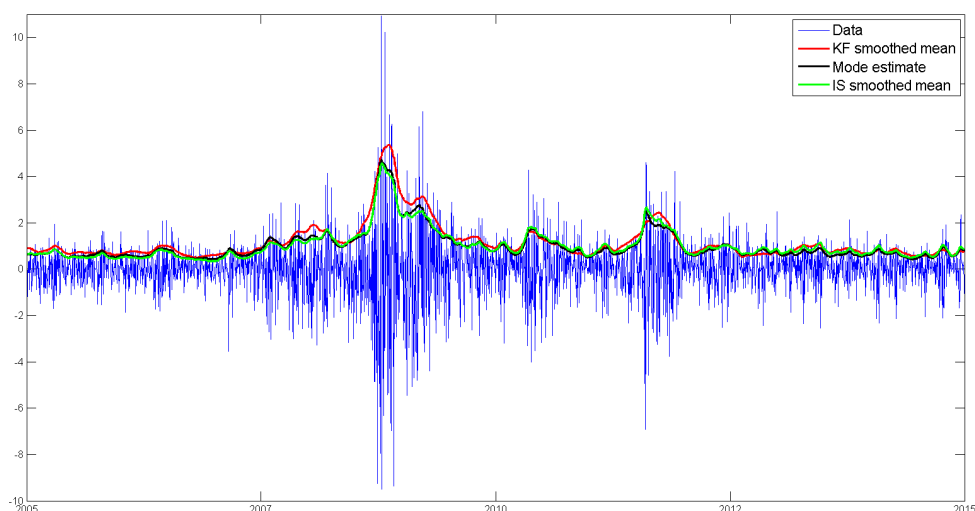
kde  $\sigma_u^2$  je rozptyl náhodné složky  $u_t$ ,  $\sigma_\eta^2$  může být interpretováno jako váha přisouzená časově specifickému šoku volatilitě, neboli volatilita volatilitě,  $\theta$  je efekt volatilitě z předcházejícího časového okamžiku na současnou hodnotu volatilitě, neboli parametr měřící autokorelaci.  $\omega$  je střední hodnota volatilitě a je součástí výrazu pro nepodmíněnou volatilitě,  $\omega/(1 - \theta)$ .

Získané odhadnuté koeficienty jsou použity pro výpočet vyhlazené odhady střední hodnoty  $h_t$  na základě aproximativního modelu pro  $x_t$  za použití metody Kalman filter smoother. Poté byly odhadnuty koeficienty pro vyhlazený modus  $h_t$ . Vyhlazené odhady střední hodnoty byl také odhadnut pomocí Importance sampling, kde počet výběrů bylo nastaveno na 100, neboli  $N = 100$ .

Výsledky pro tři různé odhady  $h_t$ , tj. odhad střední hodnoty metodou Kalman filter smoother, odhad modusu na základě Newton-Raphson algoritmu a odhad střední hodnoty pomocí importance sampling, jsou zobrazeny na obrázku 1 a 2. Hodnoty všech třech odhadů jsou si velmi blízké a všechny se zdají být rozumným odhadem volatilitě. Je však patrné, že odhad střední hodnoty z Kalman filter smoother vykazuje v průměru vyšší hodnoty než ostatní dva odhady. Je také zřejmé, že odhad vyhlazeného modusu vypočítaný na základě Newton-Raphson metody je velmi blízký vyrovnaným odhadům střední hodnoty na základě importance sampling v některých časových úsecích, avšak v průměru vykazuje o něco vyšší hodnoty.



Obrázek 1: Odhad stochastické volatilitě  $h_t$  pomocí Kalman filter smoother z aproximativního lineárního modelu, pomocí Importance sampling (IS) metody a odhad modusu  $h_t$  pomocí Newton-Raphson algoritmu.



Obrázek 2: Logaritmičké výnosy (data), vyrovnané hodnoty odhadu volatility  $\exp(h_t/2)$ , kde  $h_t$  je: vypočítáno pomocí Kalman filter smoother z aproximativního lineárního modelu; odhad modusu získaného pomocí Newton-Raphson algoritmu; odhad pomocí Importance sampling (IS) metody.

## 4 ZÁVĚR

Tato práce reportuje výsledky pro odhad volatility výnosů akcií S&P 500 pro roky 2005 až 2014. K tomuto účelu byl použit model stochastické volatility využívající Kalman filter smoother a Importance sampling. Oba přístupy k odhadu volatility poskytly podobné výsledky, přičemž odhady pomocí Kalman filter smoother jsou nepatrně vyšší než odhady za pomoci Importance sampling. Další práce se bude zabývat srovnáním odhadů z modelů stochastické volatility předpokládající  $t$ -rozdělení náhodných složek.

### Poděkování

Tento článek byl napsán v rámci grantu IGS F4/93/2017 Vysoké školy ekonomické v Praze.

### Použitá literatura

- [1] DURBIN, James; KOOPMAN, Siem Jan. Time series analysis by state space methods. Oxford University Press, 2012.
- [2] BROTO, Carmen; RUIZ, Esther. Estimation methods for stochastic volatility models: a survey. Journal of Economic Surveys, 2004, 18.5: 613-649.
- [3] GHYSELS, Eric; HARVEY, Andrew C.; RENAULT, Eric. Stochastic volatility. Handbook of statistics, 1996, 14: 119-191.
- [4] TOMANOVÁ, Petra. Analýza S&P 500 pomocí modelu stochastické volatility. In Nové trendy v ekonometrii a operačním výzkumu. Bratislava: Ekonóm, 2016, s. 162–166.

## **Kontaktní údaje**

Ing. Petra Tomanová, MSc

Vysoká škola ekonomická v Praze, Fakulta statistiky a informatiky  
nám. W. Churchilla 1938/4, 130 67 Praha 3

Tel: (420) 224 09 5445

email: [petra.tomanova@vse.cz](mailto:petra.tomanova@vse.cz)

## SCORING RULES - DEA/AR MODEL S PENALIZACÍ

## SCORING RULES - DEA/AR MODEL WITH PENALTIES

*Petra Zýková*

### **Abstrakt**

Tento článek se zabývá hlasovacími systémy s pořadím kandidátů. Cílem hlasovacího systému je určit celkového vítěze a celkové pořadí kandidátů na základě preferencí voličů. V článku jsou formulovány pokročilé hlasovací systémy s použitím DEA modelů, konkrétně DEA/AR model a DEA/AR model s penalizací. Modely jsou numericky ilustrovány na výsledcích Velkých cen Formule 1 v sezóně 2016, na Poháru konstruktérů.

***Klíčová slova:** scoring rules, DEA modely, vektor vah*

### **Abstract**

This article deals with ranked advanced voting systems. The aim is to determine the overall winner and the ranking of candidates based on voters' preferences. The article utilises advanced voting systems with the application of DEA models, specifically DEA/AR model and DEA/AR model with penalties. The models and their use are demonstrated on Formula One Grand Prix results from season 2016, on Formula One World Constructors' Championship.

***Keywords:** scoring rules, DEA models, sequence of weights*

## 1 ÚVOD

Často je třeba zjistit celkového vítěze z několika pořadí. Tímto úkolem se zabývají hlasovací systémy s pořadím kandidátů. Problém nastává při určování vah (bodů) za jednotlivá pořadí. Každý předem zvolený vektor vah může nahrávat jinému kandidátovi na celkového vítěze. Při tvorbě vektoru vah je prostor pro manipulaci výsledků. Pokud jsou již známy výsledky jednotlivých pořadí, je možné určit vektor vah pomocí DEA modelů. DEA modelem získáme pro každého kandidáta jeho optimální vektor vah, aby se necítil nebo nemohl být ošizen. Tímto se získá celkové pořadí a určí celkový vítěz. Často jsou voliči přesvědčeni, že někoho z kandidátů opravdu nechtějí. Můžou ho vyškrtnout a tím mu dát penalizaci (trestný bod). V tomto případě je možné použít DEA model s penalizací, který určí opět optimální vektor vah pro každého kandidáta.

## 2 FORMULACE ÚLOHY

Předpokládá se hlasovací systém s pořadím o  $n$  kandidátech a  $m$  voličích. Každý volič sestaví pořadí prvních  $t$  kandidátů. To znamená, že nemusí udávat celé pořadí všech kandidátů, ale pouze podle sebe určit prvních  $t$  nejlepších ( $t \leq n$ ). Volič nesmí přidělit stejné pořadí dvěma nebo více kandidátům. Cílem je určit celkového vítěze a pořadí kandidátů z jednotlivých pořadí daných voliči.

Celkové pořadí kandidátů se určí podle bodů, které  $j$ -tý kandidát obdržel tj:

$$s_j = \sum_{r=1}^t u_r y_{rj}, \text{ pro všechna } j, j = 1, \dots, n \quad (1)$$

kde  $u_r, r = 1, \dots, t$  je počet bodů získaných za  $r$ -té umístění,

$y_{rj}, r = 1, \dots, t$  a  $j = 1, \dots, n$  udává kolikrát se  $j$ -tý kandidát umístil na  $r$ -tém místě. Předpokládá se znalost  $y_{rj}$ .

Celkový vítěz je kandidát s nejvyšší hodnotou  $s_j$ .

## 3 FORMULACE DEA MODELŮ

Cílem je maximalizovat efektivnost zkoumané  $j_0$ -té jednotky. Základní DEA model lze zapsat následovně:

Maximalizovat

$$U_{j_0} = \frac{\sum_{r=1}^t u_r y_{rj_0}}{\sum_{k=1}^w v_k x_{kj_0}}, \quad (2)$$

za podmínek

$$\frac{\sum_{r=1}^t u_r y_{rj}}{r=1} \leq 1, \text{ pro } j = 1, \dots, n \quad (3)$$

$$u_r \geq \varepsilon, \text{ pro } r = 1, \dots, t, \quad (4)$$

$$v_k \geq \varepsilon, \text{ pro } k = 1, \dots, w, \quad (5)$$

kde  $\mathbf{Y} = \{y_{rj}; r = 1, \dots, t, j = 1, \dots, n\}$  je matice hodnot výstupů,

$\mathbf{X} = \{x_{kj}; k = 1, \dots, w, j = 1, \dots, n\}$  je matice hodnot vstupů,

$u_r$  je váha  $r$ -tého výstupu,

$v_k$  je váha  $k$ -tého vstupu,

$\varepsilon$  je infezimální konstanta (velmi malé číslo).

Množina přípustných řešení je určena tzv. efektivní hranicí. Produkční jednotky ležící na efektivní hranici jsou efektivní. Tj. není možná kombinace nižších vstupů se stejnými výstupy, nebo stejných vstupů s vyššími výstupy.

V případě určování pořadí kandidátů jsou kandidáti produkčními jednotkami. Každý kandidát má  $t$  výstupů – pořadí, které mu přiřadili jednotliví voliči. A pouze jednotkový vstup (DEA without explicit inputs).

DEA without explicit inputs lze zapsat následovně:

Maximalizovat

$$U_{j_0} = \sum_{r=1}^t u_r y_{rj_0}, \quad (6)$$

za podmínek

$$\sum_{r=1}^t u_r y_{rj} \leq 1, \text{ pro } j = 1, \dots, n \quad (7)$$

$$u_r \geq \varepsilon, \text{ pro } r = 1, \dots, t. \quad (8)$$

Při použití DEA modelů odpadá problém s určováním vektoru vah. Tím pádem není možné zmanipulovat vítěze. Model se řeší posupně pro všechny kandidáty  $j_0, j_0 = 1, \dots, n$ , a je založen na předpokladu, že je spravedlivé pro každého kandidáta použít vektor vah, který je pro něj optimální.

Prvním takovým modelem použitelným pro scoring rules je DEA/AR. Model lze zapsat následovně:

Maximalizovat

$$h_{j_0} = \sum_{r=1}^t u_r y_{rj_0}, \quad (9)$$

za podmínek

$$\sum_{r=1}^t u_r y_{rj} \leq 1, \text{ pro } j = 1, \dots, n, \quad (10)$$

$$u_r - u_{r+1} \geq \varepsilon, \quad r = 1, \dots, t-1, \text{ je podmínka konvexnosti vah} \quad (11)$$

$$u_t \geq \varepsilon \text{ udává velikost váhy za poslední bodované místo.} \quad (12)$$

Celkovým vítězem je kandidát s nejvyšší hodnotou  $h_{j_0}$ . Pokud  $h_{j_0} = 1$  je  $j_0$ -tý kandidát efektivní jednotkou.

#### 4 FORMULACE ÚLOHY S PENALIZACÍ

Předpokládá se hlasovací systém s pořadím o  $n$  kandidátech a  $m$  voličích. Každý volič sestaví pořadí prvních  $t$  kandidátů. To znamená, že nemusí udávat celé pořadí všech kandidátů, ale pouze podle sebe určit prvních  $t$  nejlepších ( $t \leq n$ ). Volič nesmí přidělit stejné pořadí dvou nebo více kandidátům. Dále každý volič může určit záporné kandidáty (penalizaci). Tím říká, že daného kandidáta rozhodně nechce. Cílem je určit celkového vítěze a pořadí kandidátů z jednotlivých pořadí daných voliči.

Záporné pořadí (penalizace) je označeno indexem  $t+1$ . Váha z  $t+1$  umístění je označena  $u_{t+1}$ .

Přidání penalizace rozšiřuje DEA model uvedený výše o další omezení. Prvním z rozdílů oproti DEA modelu bez penalizace je, že se účelová funkce a omezení se součtem sčítají od  $r = 1$  až po  $r = t+1$ .

Do modelu se přidává omezení

$$u_{t+1} \leq -\varepsilon, \quad (13)$$

aby záporná váha nebyla rovna nule. A dále omezení

$$\sum_{r=1}^{t+1} u_r y_{rj} \geq -1, \text{ pro } j = 1, \dots, n, \quad (14)$$

aby označení kandidáta jako záporného alespoň jedním voličem automaticky neodsoudilo kandidáta na poslední místo.

DEA/AR model s penalizací lze zapsat následovně:

Maximalizovat

$$h_{j_0}^- = \sum_{r=1}^{t+1} u_r y_{rj_0}, \quad (15)$$

za podmínek

$$\sum_{r=1}^{t+1} u_r y_{rj} \leq 1, \text{ pro } j = 1, \dots, n, \quad (16)$$

$$u_r - u_{r+1} \geq \varepsilon, \quad r = 1, \dots, t-1, \quad (17)$$

$$u_t \geq \varepsilon, \quad (18)$$

$$\sum_{r=1}^{t+1} u_r y_{rj} \geq -1, \text{ pro } j = 1, \dots, n, \quad (19)$$

$$u_{t+1} \leq -\varepsilon. \quad (20)$$

## 5 NUMERICKÁ ILUSTRACE

Jako numerická ilustrace DEA/AR modelu a DEA/AR modelu s penalizací byly použity výsledky závodů Formule 1 z roku 2016. V tabulce 1 je umístění jezdců z jednotlivých stájí ve Velkých Cenách. Každá stáj má dva jezdce. Ve sloupci technická závada je uvedeno, kolikrát jezdci dané stáje závod nedokončili z důvodu technické závady na formuli. Technická závada bude brána jako záporné hodnocení.

stáj	1	2	3	4	5	6	7	8	9	10	technická závada
Mercedes AMG Petronas	19	8	6	2	2	0	2	0	0	0	1
Red Bull Racing	2	8	6	8	4	2	4	2	0	0	2
Scuderia Ferrari	0	5	6	8	7	6	0	0	2	0	5
Sahara Force India	0	0	2	2	1	3	8	7	2	5	4
Williams Martini Racing	0	0	1	1	4	3	1	5	8	5	4
McLaren Honda Formula	0	0	0	0	2	2	4	1	4	4	6
Haas	0	0	0	0	1	1	1	1	0	1	8
Scuderia Toro Rosso	0	0	0	0	0	4	0	5	4	3	8
Renault Sport	0	0	0	0	0	0	1	0	0	2	8
Sauber	0	0	0	0	0	0	0	0	1	0	6
Manor Racing MRT	0	0	0	0	0	0	0	0	0	1	3

**Tabulka 1:** Četnost umístění stájí na jednotlivých místech a počet nedokončených závodů z důvodu technické závady, zdroj [3].



V tabulce 2 je pořadí stájí získané pomocí DEA/AR modelu. Pohár konstruktérů by získala stáj Mercedes AMG Petronas s hodnotou účelové funkce 1,0000. Druhá by skončila stáj Red Bull Racing s hodnotou účelové funkce 0,9231 a třetí stáj Scuderia Ferrari s hodnotou účelové funkce 0,8718.

<b>stáj</b>	<b>DEA/AR</b>	<b>pořadí</b>
Mercedes AMG Petronas	1,0000	1
Red Bull Racing	0,9231	2
Scuderia Ferrari	0,8718	3
Sahara Force India	0,7692	4
Williams Martini Racing	0,7179	5
McLaren Honda Formula	0,4359	6
Haas	0,1282	8
Scuderia Toro Rosso	0,4103	7
Renault Sport	0,0769	9
Sauber	0,0256	10
Manor Racing MRT	0,0256	11

**Tabulka 2:** Pořadí stájí pomocí DEA/AR modelu.

V tabulce 3 je pořadí stájí získané pomocí DEA/AR modelu s penalizací. Pohár konstruktérů by získala stáj Mercedes AMG Petronas s hodnotou účelové funkce 1,0000. Druhá by skončila stáj Red Bull Racing s hodnotou účelové funkce 0,9231 a třetí stáj Scuderia Ferrari s hodnotou účelové funkce 0,8718.

<b>stáj</b>	<b>DEAzáporná</b>	<b>pořadí</b>
Mercedes AMG Petronas	1,0000	1
Red Bull Racing	0,9231	2
Scuderia Ferrari	0,8718	3
Sahara Force India	0,7692	4
Williams Martini Racing	0,7179	5
McLaren Honda Formula	0,4359	6
Haas	0,1282	8
Scuderia Toro Rosso	0,4103	7
Renault Sport	0,0769	9
Sauber	0,0256	11
Manor Racing MRT	0,0256	10

**Tabulka 3:** Pořadí stájí pomocí DEA/AR modelu s jednou zápornou váhou.

V tabulce 4 jsou uvedeny hodnoty účelových funkcí při použití DEA/AR a DEA/AR modelu s penalizací. Neukázalo se, že zahrnutí technické závady má nějaký vliv na celkové pořadí stájí.

stáj	DEA/AR	DEA/AR záporná
Mercedes AMG Petronas	1,0000	1,0000
Red Bull Racing	0,9231	0,9231
Scuderia Ferrari	0,8718	0,8718
Sahara Force India	0,7692	0,7692
Williams Martini Racing	0,7179	0,7179
McLaren Honda Formula	0,4359	0,4359
Haas	0,1282	0,1282
Scuderia Toro Rosso	0,4103	0,4103
Renault Sport	0,0769	0,0769
Sauber	0,0256	0,0256
Manor Racing MRT	0,0256	0,0256

**Tabulka 4:** Hodnoty účelových funkcí při použití DEA/AR a DEA/AR modelu s jednou zápornou váhou.

## 6 ZÁVĚR

Určení celkového vítěze z jednotlivých pořadí představuje zajímavý úkol, jehož řešení je velmi závislé na určení vektoru vah jednotlivých pořadí. Proto jsou ve článku formulovány dva modely vycházející z metodologie analýzy obalu dat. První DEA model pracuje pouze z pořadí kandidátů daných voliči, druhý DEA model naopak bere v potaz i možnost určení nechtěného kandidáta z pohledu voliče. Navržené DEA modely byly testovány na jednoduchém numerickém příkladu – výsledcích Poháru konstruktérů Formule 1 v roce 2016. Výsledky získané z numerické ilustrace ukazují, že na zkoumaných datech neměla penalizace vliv na pořadí kandidátů. Je to dáno tím, že rozdíly v pořadí mezi kandidáty byly velmi značné, takže by byl stejný vítěz při použití jakéhokoliv způsobu určování celkového vítěze.

### **Použitá literatura**

1. HASHIMOTO, A. A ranked voting system using a DEA/AR exclusion model: A note. *European Journal of Operational Research*, 97(3): 600-604, 1997
2. JABLONSKÝ, J., DLOUHÝ M. *Modely hodnocení efektivnosti a alokace zdrojů*. Praha: Professional Publishing, 2015.
3. <http://www.vysledky.cz/automobily/formule-1/>

### **Kontaktní údaje**

Ing. Bc. Petra Zýková.

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky

Náměstí Winstona Churchilla 1938/4, 130 67 Praha 3

email: [petra.zykova@vse.cz](mailto:petra.zykova@vse.cz)