

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

30. november – 2. december 2016

30. listopad – 2. prosinec 2016

Praha

Programový výbor – recenzenti:

prof. Ing. Josef Jablonský, CSc., VŠE Praha
prof. Ing. Michal Fendek, PhD., EU Bratislava
prof. Ing. Zlatica Ivaničová, PhD., EU Bratislava
prof. RNDr. Ing. Petr Fiala, CSc., VŠE Praha
prof. Ing. Ivan Brezina, CSc., EU Bratislava
prof. RNDr. Jan Pelikán, CSc., VŠE Praha
prof. Mgr. Juraj Pekár, PhD., EU Bratislava
doc. Ing. Martin Lukáčik, PhD., EU Bratislava

Organizačný výbor:

doc. Ing. Jan Fábry, PhD.
doc. Ing. Martin Lukáčik, PhD.
Ing. Adriana Lukáčiková, PhD.
Ing. Marian Reiff, PhD.
Ing. Brian König, PhD.

Náklad: 50 ks
AH 10,36

Kontakt: jablon@vse.cz, 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 2016

ISBN: 978-80-225-4319-4

Obsah

<i>Adam Borovička:</i> New Fuzzy Weights Estimation Method	5
<i>Ivan Brezina, Juraj Pekár:</i> Vplyv hmotnosti na optimálnu trasu pri minimalizácii emisií CO₂	10
<i>Zuzana Čičková, Allan Jose Sequeira Lopez:</i> Evolučné algoritmy a riešenie úloh s ohraničeniami	15
<i>Marek Dvořák:</i> Using Artificial Neural Networks for Augmented Time Series Prediction	20
<i>Petr Fiala:</i> Zobecnění Braessova paradoxu	25
<i>Lenka Fiřtová:</i> Použití logistické regrese pro odhad parametrů položek při změnách struktury latence	30
<i>Lukáš Frýd:</i> Nelinearita v datech: je GAM model vždy králem?	39
<i>Andrea Furková, Michaela Chocholatá:</i> Softvérové nástroje na priestorovú analýzu dát	44
<i>Vladimír Holý:</i> Vliv zaokrouhlovací chyby na realizovaný rozptyl	50
<i>Jaroslav Janáček, Marek Kvet:</i> Návrh robustního záchranného systému pomocí maximálního scénáře	55
<i>Marta Janáčková, Alžbeta Szendreyová:</i> Pojednanie o efektívnom počte staníc záchranej zdravotnej služby	60
<i>Miroslava Jánošová:</i> Teoretické východiská modelu s korekčným členom	64
<i>Ivan Lichner, Marek Radvanský, Tomáš Miklošovič:</i> Efektívnosť regionálneho prerozdelenia zdrojov EÚ na úrovni NUTS 3 na Slovensku v rokoch 2007-15	68
<i>Martin Lukáčik:</i> Úvod do ekonometrickej analýzy so systémom R	77

<i>Martin Lukáčik, Adriana Lukáčiková, Karol Szomolányi:</i> Úvod do práce so systémom R	88
<i>Tomáš Majer, Stanislav Palúch:</i> Model tvorby turnusov v autobusovej doprave s rešpektovaním bezpečnostnej prestávky	99
<i>Veronika Miťková, Eduard Hozlár:</i> Modelovanie príjmov domácností v CGE modeli	105
<i>Vladimír Mlynarovič:</i> Makroekonomický vývoj a identifikácia investičných príležitostí	111
<i>Michal Pieter:</i> Nový software pre metódy triedy ELECTRE	115
<i>Tatiana Polonyankina:</i> Definice základního gravity modelu a jeho kombinace s modely prostorové autokorelace	123
<i>Marian Reiff, Zlatica Ivaničová:</i> Analýza ekonomického rastu a environmentálnych ukazovateľov pomocou klastrovej analýzy	129
<i>Allan Jose Sequeira Lopez:</i> Analýza efektívnosti podnikov v odvetví IT technológií s použitím DEA – SBM modelu a ROA	134
<i>Ondřej Sokol:</i> 1-minutová intervalová Forex data: dokážeme spočítať hornú mez rozptylu?	139
<i>Nataliya Soldatyuk:</i> Aplikace analýzy obalu dat pro měření výkonnosti zemí na letních olympijských hrách 2016	145
<i>Kvetoslava Surmanová:</i> Nezamestnanosť a hospodárske cykly	152
<i>Karol Szomolányi:</i> Sú podiely na produkcii dlhodobo konštantné?	158
<i>Petra Tomanová:</i> Analýza S&P 500 pomocí modelu stochastické volatility	162

NEW FUZZY WEIGHTS ESTIMATION METHOD

Adam Borovička

Abstract

The article deals with a new fuzzy weights estimation method. This concept is based on the linguistic expressions about the criteria relevance that are transformed to a conversation scale of the triangular fuzzy numbers. Each fuzzy number stands for a particular meaning. Then a mathematical model is specified in order to gain the concrete (quantitative) values of criteria weights. The algorithm of proposed method is applied to a determination of the criteria weights in terms of an investment decision making.

Keywords: *fuzzy number, linguistic preference, weight*

1 INTRODUCTION

If we use the multiple criteria decision making methods in order to make a decision in some situation, an expression about the criteria relevance is mostly demanded. Many methods work with quantitative information about the criteria relevance in the form of their weights. In many real cases, this information is not available. A decision maker can provide only a linguistic expression about the criteria relevance (vague preference), for instance *medium importance* or *high importance*. The proposed weights estimation method is able to transform the linguistic expressions to the form of weights by a particular conversion scale of the triangular fuzzy numbers. In order to set the final weights, a mathematical model based on a multiple objective optimization principle maximin is specified. The aim is to discuss current known approaches, identify their disadvantages and propose a new fuzzy weights estimation method whose algorithm is introduced and then applied to a determination of the criteria weights in terms of an investment decision making.

2 FUZZY WEIGHTS ESTIMATION METHOD

In this section, the basic reasons for a new fuzzy weights estimation method are specified. Then the algorithm of the proposed method is described in terms of the several steps.

2.1 Reasons for a new method

We know a several weights estimation methods based on the “defuzzification” formulae (Chen and Hwang, 1992). This procedure can be based on the optimism index (Lee, 1999), or α -level principle (Detyniecki and Yager, 2001). A determination of the value of α -level can be problematic from the site of decision maker. These approaches do not ensure weights normalization. Other concepts use a composite function in order to express the vague preferences (Narasimhan, 1981). These methods are so complicated and badly comprehensible for their users. My method eliminates the drawbacks mentioned above. The concept does not demand any additional information from a decision maker, the normalization procedure is included. The algorithm is not difficult and it is comprehensible for users. The transformation process of linguistic terms to the form of fuzzy numbers is based on the intuitive approach (see more Novák, 1989). To set the criteria weights, a mathematical model is formulated that is based on the multiple criteria optimization principle maximin (see more Bellman and Zadeh, 1970; Zimmerman, 1970). The optimization

process is inspired with known approach (Lai and Hwang, 1996). However this approach does not provide a solution (weights) in some cases, so that the necessary modifications are made.

2.2 Algorithm

The algorithm of the proposed method can be described in terms of the several following steps.

Step 1: A decision maker expresses the criteria importance linguistically from the defined scale of terms that can be formed as follows – *low*, *medium* and *high* (importance)¹. For the scale of terms, the transformation scale of triangular fuzzy numbers is proposed. The triangular fuzzy number can be formulated as $\tilde{F}_s = (w_s - \delta_s^l, w_s, w_s + \delta_s^u)$, where $s = \{low, medium, high\}$, w_s is the value describing the term s in interval $\langle 0,1 \rangle$ with the tolerance δ_s^l , or δ_s^u . As in (Novák, 1989), the transformation scale of triangular fuzzy numbers is specified via the *intuitive approach* in the following picture (Figure 1). This approach is derived from the fact that a sensation of some meaning is analogous for all people. Then the transformation scale of fuzzy numbers is designed by individual as a wide perceived concept.

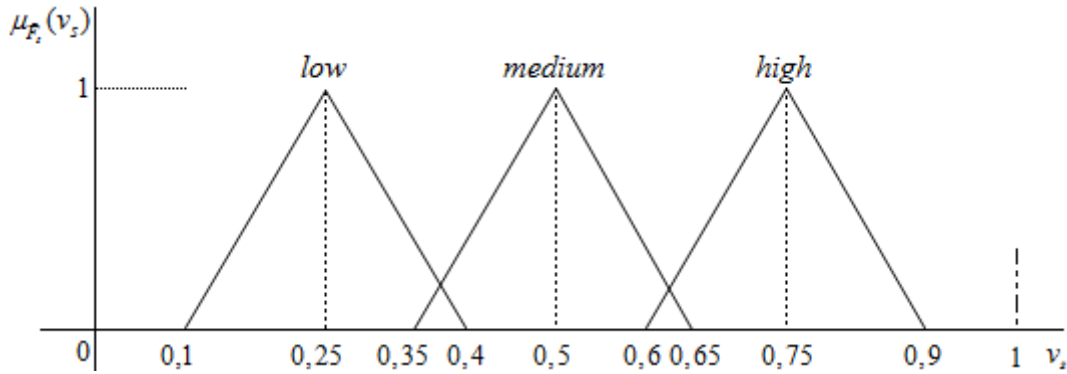


Figure 1: Transformation scale of the triangular fuzzy numbers

The membership function of fuzzy numbers must be an isosceles triangular. If it is not, a slope of the linear functions forging left parts of their membership functions must rise with a creasing linguistic term meaning, or a slope of the right linear functions forging right parts of their membership functions must decline with a creasing linguistic term meaning. Otherwise, the method provides the results against human intuition, for example the weight of a criterion with less importance can be greater than the weight of criterion with a higher importance.

Step 2: If the sum of lower limits of fuzzy numbers of the transformation scale describing expressed importance of all criteria is greater than 1, then the following modification procedure will have to be done. Otherwise, we come up to the third step. In terms of the second step, the fuzzy numbers of the transformation scale are modified as follows

$$\tilde{F}_s^{mod} = \left(\frac{w_s - \delta_s^l}{k}, \frac{w_s}{k}, \frac{w_s + \delta_s^u}{k} \right) \quad s = \{high\}$$

$$\tilde{F}_s^{mod} = \left(\frac{w_s - \delta_s^l}{K_s}, \frac{w_s}{K_s}, \frac{w_s + \delta_s^u}{K_s} \right) \quad s = \{low, medium\}$$

¹ A form of the scale can be different. The algorithm is demonstrated with the illustrative scale of 3 terms.

where

$$K_s = k \frac{1}{3} \left(\frac{w_t - \delta_t^l}{w_s - \delta_s^l} + \frac{w_t}{w_s} + \frac{w_t + \delta_t^u}{w_s + \delta_s^u} \right) \quad t = \{high\}, s = \{low, medium\}.$$

The fuzzy number expressing the highest importance is divided by the number of criteria. Other fuzzy numbers are divided by a greater value that is derived from the proportions among the values of parameters of particular triangular fuzzy number and fuzzy number expressing a highest importance. The sum of proportions is averaged. Now it is expressed the fact how many times particular fuzzy number approximately less than fuzzy number describing a highest importance. Then the fuzzy numbers are divided by a product of this value and the number of criteria. This mechanism takes into account approximately a relative position of fuzzy numbers of a transformation scale on the horizontal axis in terms of the situation of k criteria.

Step 3: If the condition of an application of the second step is not realized, the following mathematical model is solved in order to set the strict value of criteria weights

$$\begin{aligned} z = \alpha &\rightarrow \max \\ \mu_{\tilde{F}_{v_j}}(v_j) &\geq \alpha \quad j = 1, 2, \dots, k \\ \sum_{j=1}^k v_j &= 1, \\ 0 \leq \alpha &\leq 1 \\ v_j \geq 0 &\quad j = 1, 2, \dots, k \end{aligned}$$

where $\mu_{\tilde{F}_{v_j}}(v_j)$ is the membership function of a triangular fuzzy number \tilde{F}_{v_j} describing the linguistic expression of a relevance of the j -th criterion (or vague weight of the j -th criterion), v_j is the strict weight of the j -th criterion. The value α denotes a grade of membership of the solution. The model is solved via the *maximin principle*. The sum of upper limits of the fuzzy numbers of the transformation scale describing expressed importance of all criteria can be less than 1. Then a condition about the interval for the parameter α is modified to the form $\alpha \leq 1$ in the model mentioned above. Now the value α loses its original interpretation, because the grade of membership is negative. But the main idea of the model principle is kept. The value of objective function is held at the highest value as possible. Then a proportion deviation of weights from a set interval of values with the nonzero grade of membership to the defined fuzzy sets is minimized. If the modification procedure must be done in terms of the second step, then slightly different mathematical model is solved

$$\begin{aligned} z = \alpha &\rightarrow \max \\ \mu_{\tilde{F}_{v_j}^{mod}}(v_j) &\geq \alpha \quad j = 1, 2, \dots, k \\ \sum_{j=1}^k v_j &= 1, \\ \alpha &\leq 1 \\ v_j \geq 0 &\quad j = 1, 2, \dots, k \end{aligned}$$

where $\mu_{\tilde{F}_{v_j}^{mod}}(v_j)$ is the membership function of the triangular fuzzy number describing a linguistic expression about the importance of the j -th criterion in the modified form, v_j is

the strict weight of the j -th criterion. In contrast to the previous model, this modified model has some feasible solution.

3 IMPORTANCE OF CRITERIA IN THE INVESTMENT

And now some rather illustrative example of an application of the proposed method will be formulated. Imagine the investment decision making situation when a potential investor wants to invest some money in open unit trusts (see more Borovička, 2014). He/she selects a few funds from bigger group of unit trusts according to three chosen criteria – return, risk and cost. The selection is made by some multiple criteria evaluation method that demands a strict expression about the criteria relevance. However, a potential investor is able to express their relevance only linguistically. The vague preferences are formulated in the scale of 3 terms mentioned above as follows (Table 1).

Table 1: Linguistic expression of the criteria importance

Criterion	Importance
Return	<i>Medium</i>
Risk	<i>High</i>
Cost	<i>Low</i>

The mentioned above transformation scale of triangular fuzzy numbers is used (Figure 1). If the sum of lower limits of fuzzy numbers of the transformation scale describing expressed importance of all three criteria is greater than 1 ($0,1+0,35+0,6=1,1 > 1$), the modification procedure in terms of the second step of algorithm must be done. The following mathematical model is solved

$$\begin{aligned}
 z = \alpha &\rightarrow \max \\
 \frac{v_1 - 0,0761}{0,0326} &\geq \alpha & \frac{0,1413 - v_1}{0,0326} &\geq \alpha \\
 \frac{v_2 - 0,2}{0,05} &\geq \alpha & \frac{0,3 - v_2}{0,05} &\geq \alpha \\
 \frac{v_3 + 0,0089}{0,0133} &\geq \alpha & \frac{0,0355 - v_3}{0,0133} &\geq \alpha , \\
 \sum_{j=1}^3 v_j &= 1 \\
 \alpha &\leq 1 \\
 v_j &\geq 0 \quad j = 1, 2, 3
 \end{aligned}$$

where v_1 is the weight of return, v_2 denotes the weight of risk and v_3 marks the weight of cost. The final criteria weights are in the following table (Table 2).

Table 2: Final criteria weights

Criterion	Weight
Return	0,319
Risk	0,573
Cost	0,108

4 CONCLUSION

The current known fuzzy weights estimation methods are discussed and their disadvantages are identified. To eliminate them, a new weights estimation method is proposed and described in this article. The algorithm works with linguistic expressions about the criteria relevance that are transformed to the triangular fuzzy numbers. The strict weights are set by means of a mathematical model based on the multiple criteria optimization principle maximin. The proposed method is applied to a determination of the criteria relevance in terms of the investment decision making situation.

References

1. BELLMAN, R. E.; ZADEH, L. A.: Decision making in a fuzzy environment. *Management Science*. Vol. 17, No. 4, 1970, pp. 141-164.
2. BOROVIČKA, A.: New interactive multiple objective programming method applied in the investment decision making under uncertainty. *GRANT Journal*. Vol. 3, No. 1, 2014, pp. 63-67.
3. CHEN, S. J.; HWANG, C. L.: *Fuzzy Multiple Attribute Decision Making: Methods and Applications*. Berlin: Springer-Verlag, 1992, ISBN 3-540-54998-6.
4. DETYNIĘCKI, M.; YAGER, R. R.: Ranking fuzzy numbers using α -weighted valuations. *Fuzziness and Knowledge Based Systems*. Vol. 8, No. 5, 2001, pp. 573-592.
5. LAI, Y. J.; HWANG, C. L.: *Fuzzy Multiple Objective Decision Making: Methods and Applications*. 2nd ed. Berlin: Springer, 1996, ISBN 3-540-57595-2.
6. LEE, A. R.: Application of Modified Fuzzy AHP Method to Analyze Bolting Sequence to Structural Joints. *Ph.D. Thesis*. USA: Bell & Howell Company, 1999.
7. NARASIMHAN, R.: On fuzzy goal programming - some comments. *Decision Science*. Vol. 12, No. 3, 1981, pp. 532-538.
8. NOVÁK, V.: *Fuzzy Sets and Their Applications*. Bristol: Adam Hilger, 1989, ISBN 978-0852745830.
9. ZIMMERMAN, H. J.: Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems*. Vol. 1, No. 1, 1970, pp. 45-55.

Contact

University of Economics in Prague, Faculty of Informatics and Statistics, Department of Econometrics

W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic

E-mail: adam.borovicka@vse.cz

Telephone number: +420 224 09 5447

The research project was supported by Grant No. IGA F4/54/2015 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

VPLYV HMOTNOSTI NA OPTIMÁLNU TRASU PRI MINIMALIZÁCII EMISIÍ CO₂¹

IMPLICATIONS FOR WEIGHT OPTIMAL ROUTE FOR MINIMIZING CO₂ EMISSIONS

Ivan Brezina, Juraj Pekár

Abstrakt

V súčasnosti predstavuje zaujímavú „zelenú“ modifikáciu klasických úloh CVRP minimalizácia emisií CO₂, ktorú reprezentujú modely GVRP (*Green Vehicle Routing Problem*). V príspevku je cieľom modifikácie klasických CVRP minimalizácia emisií CO₂ vozidiel stanovením poradia navštívených odberných miest. Cieľom je nájsť optimálny rozvoz, resp. zvoz tovaru, pričom v modeli predpokladáme, že na ich tvorbu má vplyv zaťaženie vozidla.

Kľúčové slová: CVRP, GVRP

Abstract

The paper presents an interesting "green" modification of the CVRP minimization of emissions of CO₂, which represent the models GVRP (*Green Vehicle Routing Problem*). In this article the aim is to minimize modifications of classical CVRP CO₂ emission vehicles specifying the order of sampling sites visited. The aim is to find optimal distribution of goods, in which the model assumes that their work has an impact load.

Keywords: CVRP, GVRP

1 ÚVOD

V príspevku budeme vychádzať z klasickej úlohy rozvozu tovaru s kapacitným obmedzením vozidla. Túto úlohu možno charakterizovať aj ako úlohu obchodného cestujúceho so špeciálnou štruktúrou (Čičková, 2013), respektíve je takáto úloha v literatúre označovaná ako úloha okružných jázd (*Vehicle Routing Problem – VRP*), resp. častejšie ako kapacitná okružná úloha (*Capacited Vehicle Routing Problem – CVRP*).

Jednou z možných modifikácií klasickej CVRP môže byť jej rozšírenie o minimalizáciu škodlivých plodín, predovšetkým emisií CO₂. V takomto prípade optimalizácie okružných rozvozných trás s cieľom minimalizácie emisií CO₂ hovoríme o emisnej úlohe okružných jázd (*Emissions Vehicle Routing Problem, EVRP*). V takejto modifikácii je optimalizácia (minimalizácia) prejdenej vzdialenosti vozidlom nahradená cieľovým kritériom minimalizácie vyprodukovaných emisií CO₂ (pri predpoklade závislosti spotreby paliva na hmotnosti vozidla v priebehu jeho cesty).

2 ÚLOHA ROZVOZU TOVARU

Uvažujme ďalej o preprave objemnejších komodít z centra zákazníkom prostredníctvom nákladného vozidla produkujúceho emisie CO₂ v závislosti od vlastnej hmotnosti vozidla

¹ This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. 1/0245/15 „Transportation planning focused on greenhouse gases emission reduction“.

a od hmotnosti prepravovaného tovaru. Každému zákazníkovi je pritom priradená požiadavka na obsluhu (požadované množstvo komodity). Pri plánovaní prepravy je pritom nevyhnutné zohľadniť aj obmedzenie kapacity prepravného prostriedku. Tento problém je z literatúry známy ako problém rozvozu tovaru (*Vehicle Routing Problem – VRP*), resp. kapacitný okružný problém (*Capacited Vehicle Routing Problem – CVRP*). Jeho podstatou je nájsť také minimálne trasy vozidla, ktoré sa začínajú a končia vo východiskovom (aj koncovom) uzle tak, aby každý uzol (zákazník) okrem východiskového uzla (centra) bol navštívený práve raz, pričom sa pri preprave neprekročí kapacita vozidla a budú uspokojené požiadavky všetkých zákazníkov.

Označenie:

$N = \{1, 2, \dots, n\}$ – množina odberateľských uzlov (zákazníkov),

$N_0 = N \cup \{0\}$ – množina všetkých uzlov (odberateľov spolu s centrom),

d_{ij} – minimálna vzdialenosť medzi uzlami $i, j \in N_0$,

q_i – požiadavky uzla (odberateľa) $i \in N$,

g – kapacita vozidla.

Úlohu CVRP potom možno zapísať nasledujúcim spôsobom:

$$\min f(\mathbf{X}, \mathbf{u}) = \sum_{i \in N_0} \sum_{j \in N_0} d_{ij} x_{ij} \quad (1)$$

$$\sum_{i \in N_0} x_{ij} = 1, \quad j \in N, \quad i \neq j \quad (2)$$

$$\sum_{j \in N_0} x_{ij} = 1, \quad i \in N, \quad i \neq j \quad (3)$$

$$u_i + q_j - g(1 - x_{ij}) \leq u_j, \quad i \in N_0, \quad j \in N, \quad i \neq j \quad (4)$$

$$q_i \leq u_i \leq g, \quad j \in N \quad (5)$$

$$u_0 = 0 \quad (6)$$

$$x_{ij} \in \{0, 1\}, \quad i, j \in N_0, \quad i \neq j \quad (7)$$

3 EMISNÉ ÚLOHY OKRUŽNÝCH JÁZD

Široké možnosti zmien ohraničení úlohy rozvozu/zvozu dokážu do modelu zahrnúť aj aspekt ekológie v podobe optimalizácie emisných nárokov na prepravu. Aplikácie tohto problému sa objavujú už od roku 2006. Tieto modely počítajú aj s parametrami ako efektívnosť vozidiel, prepravná rýchlosť atď. Úlohy Green Vehicle Routing Problem – GVRP môžeme rozdeliť do troch hlavných skupín:

1. **Green-VRP** – väčšina z nich sa zameriava na optimalizáciu spotreby energie vozidiel v podobe palivových nárokov. Z publikovaných modelov vyberáme nasledovné modely: **EMVRP** - Úloha VRP minimalizujúca energetické nároky vozidla², **FCVRP** – Úloha CVRP s parametrom miery spotreby.
2. **Pollution Routing Problem (PRP)** – modely zamerané priamo na emisie, pričom účelová funkcia týchto modelov často naberá na zložitosti. Patria sem modely ako

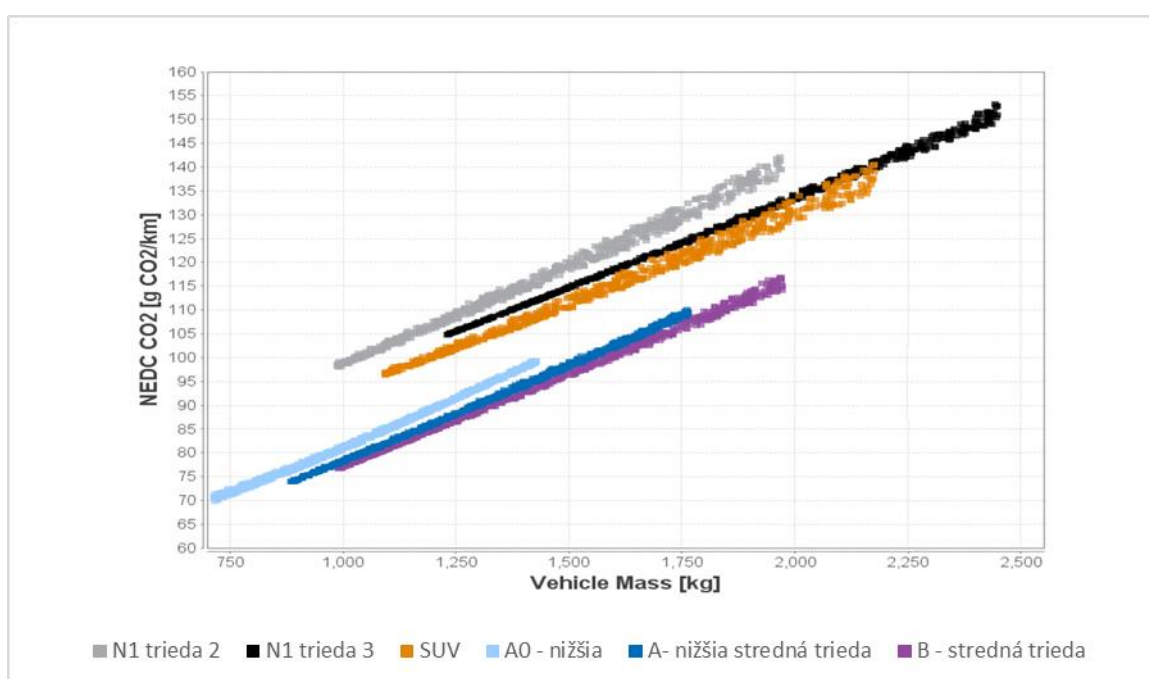
² KARA, I. – KARA, B. – YETIS, M.: *Combinatorial Optimization and Applications: Energy Minimizing Vehicle Routing Problem*. 2007 p. 62-71

TDVRP – Časovo závislý VRP, ktorý sa snaží optimalizovať objem emisií prostredníctvom skrátenia doby prepravy pri optimálnej prepravnej rýchlosti³, **EVRP-VC** - Emisná úloha rozvozu s heterogénnym dopravným parkom

- Úlohy VRP v reverznej logistike (VRPRL)** – predstavujú úlohy reverznej logistiky zamerané na aspekt rozvozu/zvozu. Patria sem úlohy z oblasti zberu odpadu, simultánny rozvoz a zvoz alebo modely selektívneho zberu s ocenením.⁴

4 TVORBA MODELU EVRP

V tejto časti vytvoríme optimalizačný model emisií s prihliadnutím na zmenu hmotnosti vozidla pri rozvoze (zvoze) tovaru. Vychádzame pritom z logického predpokladu zníženia celkovej hmotnosti použitého vozidla pri uspokojení požiadavky konkrétneho uzla (odberateľa). Zmena váhy vozidla taktiež vplýva na palivové nároky a teda aj na objem vyprodukovaných emisií. Na grafe č. 1 vidíme vzájomnú regresiu medzi objemom CO₂ na kilometer jazdy a hmotnosťou pre rôzne kategórie vozidiel.



Graf 1: Vplyv hmotnosti vozidla na výšku celkových emisií

Zdroj: KASAB, J.J. - VELLIYIUR. 2012. *Analysis of Greenhouse Gas Emission Reduction Potential of Light Duty Vehicle Technologies in the European Union for 2020–2025* (výskumná správa). Dostupné na internete: <http://www.theicct.org/sites/default/files/publications/Ricardo_LDVEU%20Technology%20Potential%20Analysis.pdf> s. 41

Pridaním parametra hmotnosti prepravovaného tovaru do základnej úlohy CVRP a pridaním parametrov ovplyvňujúcich vyprodukovaný objem emisií CO₂ získavame model optimalizujúci emisie pri rozvoze (zvoze) tovaru.

Označenie:

$N = \{1, 2, \dots, n\}$ – množina uvažovaných odberateľských uzlov (odberateľov),

³ FAGERHOLT, K. - LAPORTE, G. - NORSTAD, I.: Reducing fuel emissions by optimizing speed on shipping routes. 2010. In: *Journal of the Operational Research Society* 61, p. 523-529.

⁴ CANHONG, L. et al.: Survey of Green Vehicle Routing Problem: Past and future trends. 2013 In: *Expert Systems with Applications: An International Journal* 1127- 1132

$N_0 = \{0, 1, \dots, n\}$ – množina všetkých uvažovaných uzlov (odberateľov vrátane centra),

$M = \{1, 2, \dots, m\}$ – množina uvažovaných vozidiel,

a – produkcia emisií CO₂ vozidla na 1 km,

b – dodatočná produkcia emisií CO₂ vozidla na kg užitočného zaťaženia na 1 km,

Matematický model tejto úlohy má pri dodržaní predchádzajúceho označenia nasledujúci tvar:

$$\min f(\mathbf{X}, \mathbf{u}) = \sum_{i \in N_0} \sum_{j \in N} (d_{ij} (ax_{ij} + bu_i x_{ij})) \quad (8)$$

$$\sum_{i \in N_0} x_{ij} = 1, \quad j \in N, \quad i \neq j \quad (9)$$

$$\sum_{j \in N_0} x_{ij} = 1, \quad i \in N, \quad i \neq j \quad (10)$$

$$u_i + q_j - g(1 - x_{ij}) \leq u_j, \quad i \in N_0, \quad j \in N, \quad i \neq j \quad (11)$$

$$q_i \leq u_i \leq g, \quad j \in N \quad (12)$$

$$u_0 = 0 \quad (13)$$

$$x_{ij} \in \{0, 1\}, \quad i, j \in N_0, \quad i \neq j \quad (14)$$

V kombinácii s množstvom naloženého tovaru vo vozidle v jednotlivých uzloch (u_i) možno nájsť optimálnu cestu, pri ktorej je celkový objem vyprodukovaných emisií minimálny. Model zároveň reaguje na hmotnosť vozidla zmenou hodnoty na základe parametra b pri vyložení (naložení) tovaru v konkrétnom uzle. Takáto úprava ale značne zvyšuje výpočtovú náročnosť modelu, keďže úloha má nelineárny charakter.

Záver

Prezentovaný model je modifikáciou klasického CVRP, pričom jeho cieľom je znižovanie emisií CO₂. V tejto modifikácii je optimalizácia prejdenej vzdialenosti vozidlom nahradená cieľovým kritériom minimalizácie emisií (pri predpoklade závislosti spotreby paliva na hmotnosti vozidla v priebehu jeho cesty). V príspevku je uvedená matematická formulácia uvedeného modelu, pričom problémom je jeho riešiteľnosť, keďže uvedený model je nelineárny.

Použitá literatúra

1. CANHONG, L. et al.: Survey of Green Vehicle Routing Problem: Past and future trends. 2013 In: Expert Systems with Applications: An International Journal, p. 1127- 1132
2. ČIČKOVÁ, Z. 2008. *Optimalizácia rozvozu materiálu*. In: Medzinárodný seminár mladých vedeckých pracovníkov. [online]. Bratislava: Vydavateľstvo Ekonóm, 2008. ISBN 978-80-225-2498-8.
3. ČIČKOVÁ, Z. 2013. Modifikácie problému obchodného cestujúceho / Zuzana Čičková. - VEGA 1/0104/12. In AIESA – budovanie spoločnosti založenej na vedomostiach [elektronický zdroj] : zborník : 15. medzinárodná vedecká konferencia : Bratislava, 25. október 2013 = AIESA - building of society based on knowledge : proceedings / zostavenie zborníka Anna Šlosárová, Janette Brixová ; recenzenti Igor Bandurič, Ivan Brezina. - Bratislava : Vydavateľstvo EKONÓM, 2013. - ISBN 978-80-225-3770-4. - S. 1-11 CD-ROM.

4. FAGERHOLT, K. - LAPORTE, G. - NORSTAD, I.: Reducing fuel emissions by optimizing speed on shipping routes. 2010. In: *Journal of the Operational Research Society* 61, p. 523-529.
5. KARA, I. – KARA, B. – YETIS, M. 2007. *Energy minimizing vehicle routing problem*. In: *Combinatorial optimization and application* (pp. 62–71). [online]. 2007, vol. 4616 [cit. 2015.02.20]. Dostupné: <http://link.springer.com/chapter/10.1007%2F978-3-540-73556-4_9>.
6. KASAB, J.J. - VELLIYIUR. 2012. *Analysis of Greenhouse Gas Emission Reduction Potential of Light Duty Vehicle Technologies in the European Union for 2020–2025* (výskumná správa). Dostupné na internete: <http://www.theicct.org/sites/default/files/publications/Ricardo_LDV%20EU%20Technology%20Potential%20Analysis.pdf> s. 41
7. KOPFER, W. – SCHÖNBERGER, J. – KOPFER, H. 2013. *Reducing greenhouse gas emissions of a heterogeneous vehicle fleet*. In: *Flexible Services and Manufacturing Journal* [online]. 2013 [cit. 2015.03.13]. Dostupné: <<http://link.springer.com/article/10.1007%2Fs10696-013-9180-9>>.ISSN 1936-6582.
8. SCOTT, C. – URQUHART, N. – HART, E. 2010. *Influence of topology and payload on CO2 optimized vehicle routing*. In: *Proceedings of the 2010 international conference on applications of evolutionary computation* [online]. Vol. 2, p. 141–150. 2010. [cit. 2015.04.10]. Dostupné: <http://link.springer.com/chapter/10.1007/978-3-642-12242-2_15>. ISSN 0302-9743.

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, mail: ivan.brezina@euba.sk

prof. Mgr. Juraj Pekár, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava; Tel: (421 2) 67 295 829, mail: juraj.pekar@euba.sk

EVOLUČNÉ ALGORITMY A RIEŠENIE ÚLOH S OHRANIČENIAMÍ

SOLVING CONSTRAINED PROBLEMS BY EVOLUTIONARY ALGORITHMS¹

Zuzana Čičková, Allan Jose Sequeira Lopez

Abstrakt

Evolučné algoritmy sú výpočtové techniky, ktoré možno použiť na riešenie širokého spektra rôznych úloh. Riešenie úloh s obmedzeniami však vyžaduje určitú modifikáciu základného typu príslušného algoritmu, pretože aplikácia základných prístupov zvyčajne vedie k vzniku neprípustných riešení. V príspevku prezentujeme rôzne prístupy, ktoré možno použiť na riešenie úloh s obmedzeniami.

KLúčové slová: evolučné algoritmy, úlohy s obmedzeniami

Abstract

Evolutionary algorithms belong to computational techniques that can be used to address a wide range of different problems. However, solving problems with constraints usually requires some modification of the basic type of the algorithm, since the application of basic approaches usually leads to unfeasible solutions. In this paper we present different approaches that can be used to solve constrained problems with restrictions.

Keywords: evolutionary algorithms, constrained problems

1 EVOLUČNÉ ALGORITMY

Evolučné algoritmy sú algoritmy založené na princípoch spadajúcich pod pojem spätné inžinierstvo (Kvasnička, 2000), čo znamená, že tieto princípy sú odpozorované z prírody (najviac využívaný je princíp evolúcie), alebo z iných vedných disciplín (napr. žíhanie tuhých telies vo fyzike). Okrem genetických algoritmov je známych viacero príbuzných evolučných techník ako napr. mravčí algoritmus (Ant Colony Optimization) (Zelinka, 2002), (Brezina, 2011), metóda imunologického systému (Immunology System Method) (Lampinen, 1999), memetické algoritmy (Memetic Algorithm) (Onwubolu, 2004), rozptylové prehľadávanie (Scatter Search) (Onwubolu, 2004), rojenie častíc (Particle Swarm) (Zelinka, 2002), (Onwubolu, 2004), diferenciálna evolúcia (Differential Evolution) (Zelinka, 2002), (Onwubolu, 2004), (Čičková, 2005), samoorganizujúci sa migračný algoritmus (Self Organizing Migrating Algorithm) (Zelinka, 2002), (Onwubolu, 2004), (Čičková, 2005) a pod. Terminológia vlastná evolučným algoritmom vychádza z biologickej analógie. Tradičné optimalizačné metódy vychádzajú zo začiatočného vhodného (alebo náhodne zvoleného) odhadu riešenia, ktoré iteratívne vylepšujú. Na rozdiel od klasických techník, evolučné algoritmy patria do množiny tzv. populačných prehľadávacích algoritmov, t. j. pracujú súčasne s množinou riešení. Ak využijeme terminológiu vlastnú evolučným technikám, sú založené na práci s populáciou jedincov, kde každý jedinec (niekedy tiež používame pojem chromozóm) predstavuje aktuálne riešenie daného problému (ďalej budeme používať pojmy

¹ This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. 1/0245/15 „Transportation planning focused on greenhouse gases emission reduction“.

riešenie a jedinec v rovnakom význame). V podstate je to množina argumentov účelovej funkcie, ktorej optimálnu hodnotu hľadáme. Z hľadiska fungovania algoritmov je dôležitý postupný vývoj, kooperácia a fungovanie celej populácie (vývoj populácie prebieha spravidla v diskretných krokoch). Evolučné algoritmy vo všeobecnosti dobre fungujú pri prehľadávaní celého priestoru možných riešení. Riešenie úloh s obmedzeniami však vyžaduje určité modifikácie základných typov algoritmov, pretože všeobecné princípy by viedli k vzniku neprípustných jedincov.

2 RIEŠENIE ÚLOH S OBMEDZENIAMÍ

Na riešenie úloh s ohraničeniami možno vo všeobecnosti využiť rôzne prístupy, ktoré sú napr. v (Mařík, 2001) rozdelené na 3 skupiny: pri použití penalizačných funkcií je ohodnotenie neprípustných jedincov penalizované vhodne zvoleným penalizačným vzťahom, pri použití dekodérov a opravných algoritmov sú neprípustní jedinci transformovaní tak, aby charakterizovali prípustné riešenia a reprezentácia závislá od typu úlohy spojená so špeciálnymi rekombinačnými operátormi umožňuje vznik len prípustných jedincov. Vo (Venkatraman, 2005) sú uvedené odlišné prístupy k riešeniu úloh s ohraničujúcimi podmienkami. Penalizačné funkcie umožňujú okrem znevýhodnenia neprípustného jedinca aj použitie bariérových funkcií (tzv. death penalizácia), metódy založené na viackriteriálnej optimalizácii fungujú na dvojakom posudzovaní jedincov (na základe ich ohodnotenia a tiež na základe ich prípustnosti) a tretím prístupom je preferencia prípustného riešenia pred neprípustným riešením, ktorý zohľadňuje, že ohodnotenie neprípustných jedincov je vždy menšie ako ohodnotenie jedincov prípustných. Iný spôsob delenia známych metód uvádzajú Craenen, Eiben a Hemert v (Craenen, 2003). Prístupy rozdeľujú na tzv. prístupy s priamym zohľadnením obmedzení, keď porušenie ohraničení nie je zahrnuté priamo v účelovej funkcii (napr. eliminovanie neprípustných jedincov, oprava neprípustných jedincov, použitie špeciálnych rekombinačných operátorov, dekódovanie (transformácia) priestoru riešení) a na tzv. prístupy s nepriamym zohľadnením obmedzení (penalizácia porušenia prípustnosti z hľadiska ohraničení a penalizácia porušenia prípustnosti z hľadiska premenných).

Aj keď pohľady na spôsob riešenia úloh s obmedzeniami použitím evolučných algoritmov sú rôzne, ďalej uvedieme stručnú charakteristiku niektorých prístupov.

Penalizačné funkcie. Pri tomto prístupe sú jedinci, ktorí predstavujú neprípustné riešenia, zámerne znevýhodňovaní a evolučný princíp by mal zabezpečiť, že takíto jedinci sú postupne z populácie eliminovaní. Penalizáciu možno však použiť len v prípade, že sa dá pre tieto neprípustné riešenia vyčíslieť ich ohodnotenie. Najjednoduchší spôsob penalizácie je použitie vhodne zvoleného penalizačného člena, ktorý je prirátaný k ohodnoteniu neprípustného jedinca (v prípade minimalizácie funkcie). Známe sú však aj mnohé iné tvary penalizačných funkcií (napr. Zelinka, 2002). Ďalší možný spôsob penalizácie je stanovenie veľkosti penalizačného člena v závislosti od počtu porušených ohraničení a od „veľkosti“ ich porušenia, čo jednoduchým spôsobom zaručuje, že „veľkosť neprípustnosti“ jedinca je úmerná jeho znevýhodneniu. Takýto spôsob je uvedený napr. v (Hamida, 2000) pod názvom statická penalizácia. Na rozdiel od statickej penalizácie je tzv. dynamická penalizácia založená na myšlienke zvyšovania penalizačného zaťaženia počas evolúcie (penalizácia je tým väčšia, čím vyššia je iterácia, v ktorej sa neprípustné riešenie vyskytlo). Možným typom dynamickej penalizácie je tzv. chladiaca penalizácia (Michalewicz, 1994), kde aktuálna penalizácia závisí od tzv. chladiacej teploty, ktorá sa zvyšuje počas každej iterácie. Adaptívna penalizácia (Smith, 1993) je prístupom, keď penalizácia závisí od ohodnotenia najlepšieho jedinca v populácii. Uvedené typy penalizačných prístupov však vyžadujú nastavenie rôznych parametrov, čo vo veľkej miere sťažuje ich aplikovateľnosť. Tento nedostatok sa snažili

odstrániť napr. Farmani a Wright (Farmani, 2003) pomocou tzv. samoadaptívnej penalizačnej funkcie, ktorá vyžaduje definovanie len jedného faktora, pričom je zabezpečené, že najhorší jedinec je penalizovaný mierou, ktorá je väčšia alebo sa rovná ohodnoteniu najlepšieho jedinca v populácii a ostatní neprípustní jedinci sú penalizovaní menšou mierou v závislosti od ich neprípustnosti. Pri viacerých druhoch penalizácie je potom teda pravdepodobné, že „málo neprípustný“ jedinec má lepšie ohodnotenie ako „zlý“ prípustný jedinec.

Bariérové funkcie (niekedy označované ako „death“ penalizácia (Venkatraman, 2005), (Hamida, 2000). V takomto prípade sú neprípustné riešenia jednoznačne zakázané, tzn., že takéto riešenia neovplyvňujú populáciu v nasledujúcej iterácii. Takto však možno dosiahnuť taký stav, keď sa v celej populácii nachádzajú len neprípustní jedinci a je nutné vygenerovať novú populáciu. Aj keď je použitie bariérových funkcií ľahko implementovateľné, v praxi sa neodporúča, pretože môže viesť k „rozbitiu“ priestoru možných riešení a toto zníženie diverzibility populácie negatívne obmedzí priebeh evolučného procesu. Vo všeobecnosti možno povedať, že použitie bariérových funkcií je efektívne len v prípade konvexnej množiny prípustných riešení.

Metódy založené na preferencii prípustných jedincov. Tento prístup je podobný ako prístup penalizačný, s tým rozdielom, že zabezpečí väčšie ohodnotenie aj pre najhoršieho prípustného jedinca v porovnaní s jedincami neprípustnými. Zvyčajne sú potom prípustné jedince porovnávané na základe ich ohodnotenia, kým neprípustné jedince sú porovnávané na základe toho, v akej miere porušujú ohraničenia úlohy.

Dekodéry a opravné algoritmy. Sú to špeciálne procedúry, ktoré sa snažia prekódovať neprípustných jedincov tak, aby reprezentovali prípustné riešenia úlohy. Dekódovanie prebieha tak, že sa nájde (v zmysle použitej reprezentácie), blízky prípustný jedinec, ktorého ohodnotenie potom považujeme za ohodnotenie neprípustného jedinca. Návrh dekodéra však býva zvyčajne veľmi zložitý. Ak práve skúmaného jedinca nahradíme jeho „blízkym“ jedincom, použitý postup nazývame opravou. Opravné algoritmy však môžu zaviesť populáciu do lokálneho extrému, ak však opravu nepoužijeme, procedúra výpočtu algoritmu bude vyžadovať časté dekódovanie, čo bude mať za následok vysokú výpočtovú náročnosť algoritmu (Mařík, 2001).

Reprezentácia závislá od typu úlohy spojená so špeciálnymi rekombinačnými operátormi. Takýto postup sa ukazuje ako efektívny, pretože je založený na špeciálnom kódovaní jedincov a špeciálnej rekombinácii, ktoré eliminujú neprípustné riešenia. Súčasne však kladie vysoké nároky na formuláciu úlohy a vyžaduje dôkladnú znalosť skúmanej problematiky, pretože treba vytvárať nové operátory zvlášť pre každý typ úlohy. Zaujímavý spôsob použitia rekombinačných operátorov je uvedený v (Schoenauer, 1996), kde hlavnou ideou je myšlienka, že v mnohých prípadoch sa globálny extrém nachádza na hranici množiny prípustných riešení. Preto sa ako prvé vymedzí priestor možných riešení a následne sú použité špeciálne rekombinačné operátory, ktoré zabezpečia preskúmanie jeho hraníc.

Metódy založené na viackriteriálnej optimalizácii. Hlavným princípom je uvažovať oddelene jedincov a obmedzenia, kde napr. zoradíme najskôr riešenia na základe ich prípustnosti a až potom sa zohľadňuje účelová funkcia. V (Coello, 2000), je uvedený prístup použitia viackriteriálnej optimalizácie použitím subpopulácií, ktorých počet je rovnaký ako počet ohraničení zvýšený o jednu subpopuláciu, ktorá je posudzovaná len na základe účelovej funkcie. Pre ostatné subpopulácie sa určí vhodnosť na základe porušenia príslušných obmedzení. V každej takejto subpopulácii sa potom minimalizuje celková neprípustnosť.

Z uvedeného vyplýva, že veľké výhody, ktoré ponúka použitie evolučných techník môžu byť značne redukované nevhodným zvolením metódy, ktorá zabezpečí prípustnosť jednotlivých riešení. Všeobecne použiteľné metódy zvyčajne tolerujú prítomnosť neprípustných jedincov v

populácii, čo môže radikálnym spôsobom ovplyvniť rýchlosť výpočtu, pretože niekedy iteratívne náročné vytváranie nových neprípustných jedincov nemusí vnieť do celého procesu užitočnú informáciu, čím sa celý proces zbytočne spomaľuje. Použitie takýchto metód súčasne vyžaduje apriórne nastavenie parametrov, ku ktorému neexistuje hlbšia teoretická základňa. Použitie špecializovaných metód však môže obmedzovať ich zložitý návrh a nutnosť detailnej znalosti daného problému. Rozdiel v efektívnosti prístupov je dobre viditeľný napr. z práce (Čičková, 2005), kde je porovnaný penalizačný prístup na riešenie úlohy obchodného cestujúceho algoritmom diferenciálnej evolúcie a samoorganizujúcim sa migračným algoritmom a prístup založený na transformácii neprípustných riešení na riešenia prípustné.

Použitá literatúra

1. BREZINA, I. - ČIČKOVÁ, Z. 2011. Solving the travelling salesman problem using the ant colony optimization. *Management information systems*, vol. 6, no. 4, pg 10-14. 2011.
2. COELLO, C. A. 2000. Treating Constraints as Objectives for Single-Objective Evolutionary Optimization. *Engineering Optimization*, vol. 32, no. 3, pg. 275-308. 2000.
3. CRAENEN, B. G. W. – EIBEN, A. E. – HEMERT, J. I. 2003. *Comparing Evolutionary Algorithms on Binary Constraint Satisfaction Problems*. s.l. : IEEE Transaction on Evolutionary Computation, vol. 7, no. 5, 2003.
4. ČIČKOVÁ, Z. 2005. Aplikácia evolučných algoritmov na riešenie úlohy obchodného cestujúceho. *Dizertačná práca*. Bratislava : Katedra operačného výskumu a ekonometrie Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave, 2005.
5. FARMANI, R. – WRIGHT, J. 2003. Self-Adaptive fitness formulation for constrained optimization. *IEEE Transaction on Evolutionary Computation*, vol. 7, no. 5, pg. 445-455. 2003.
6. HAMIDA, S., B. – PETROWSKI, A. 2000. The need for improving the exploration operators for constrained optimization problems. *Proceedings of the IEEE Congress on Evolutionary Computation*, pg. 1176-83. Los Alamitos : IEEE Press, 2000.
7. KVASNIČKA, V. - POSPÍCHAL, J. - TIŇO, P. 2000. *Evolučné algoritmy*. Bratislava : STU, 2000. ISBN 80-227-1377-5.
8. LAMPINEN, J. - ZELINKA, I. 1999. *Mechanical Engineering Design Optimisation by Differential Evolution*. London : McGraw-Hill, 1999. ISBN 007-709506-5.
9. MAŘÍK, V. - ŠTĚPÁNKOVÁ, O. - LAŽANSKÝ, J. 2001. *Umělá inteligence 3*. Praha : Academia Praha, 2001. ISBN 80-200-0472-6.
10. MICHALEWICZ, Z. – ATTIA, N. 1994. Evolutionary optimization of constrained problems. *Proceeding of the 3-rd Annual Conference on Evolutionary Programming*. 1994.
11. ONWUBOLU, G. C. - BABU, B. V. 2004. *New Optimization Techniques in Engineering, Studies in Fuzziness and Soft Computing*. Berlin, Heidelberg, New York, Hong Kong, London, Milano, Paris, Tokyo : Springer-Verlag, 2004. ISBN 978-3-540-39930-8.
12. SCHOENAUER, M. – MICHALEWICZ, Z. 1996. *Evolutionary computation at the edge of feasibility*. Berlin : Springer Verlag, 1996. ISBN 978-3-540-61723-5.

13. SMITH, A. – TATE, D. 1993. Genetic optimization using a penalty functions. *Proceeding of the 5-th International Conference on Genetic Algorithms*. s.l. : Morgan Kaufmann, 1993.
14. VENKATRAMAN, S. – YEN, G., G. 2005. *A Generic Framework for Constrained Optimization Using Genetic Algorithms*. s.l. : IEEE Transaction on Evolutionary Computation, vol. 9, no. 4, 2005.
15. ZELINKA, I. 2002. *Umělá intelingence v problémech globální optimalizace*. s.l. : BEN, 2002. ISBN 80-7300-069-5.

Kontaktné údaje

Doc. Ing. Zuzana Čičková, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemská 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 828

email: cickova@euba.sk

USING ARTIFICIAL NEURAL NETWORKS FOR AUGMENTED TIME SERIES PREDICTION

*Marek Dvořák*¹

Abstract Artificial neural networks are a powerful tool. In this paper, I am introducing a stock market time series prediction method based on data augmentation coupled with neural network approach. Introduction of sliding windows allows me to calculate several dozens of small statistics across this sliding window forming the basis for explanatory variables.

A BSI index, calculated from yet another sliding window, tells us whether a stock purchase would be beneficial and acts as a dependent variable.

Keywords: neural net, data extension, time series, prediction

1 INTRODUCTION

Classic time series predicting models rely most predominantly on some variant of autoregression models like ARIMA or VAR. These are old methods thoroughly explored and described. Extensions to those include for example seasonal adjustments which aim to decrease some variability in predictions. VAR models also use multidimensional input data to estimate lagged interdependency on each other.

Model described in this paper does not take multi-dimensionality into consideration and instead focuses on prediction based on a single time series. From this time series, a so called BSI (Buy/Sell Index) is computed and later predicted by the artificial neural network. BSI is a binary indicator telling us if we should invest in a particular stock.

2 ARTIFICIAL NEURAL NETWORK

Artificial neural net takes inspiration from biological neurons which can be found in central neural systems of humans and many other animals. Such net connects units called neurons, that are defined by its inputs, output and body of the neuron itself. [2] Such neuron can be simply thought of as a function which takes many inputs, weights them, and outputs a single value filtered through an activation function. This output can be used as an input to another neuron. Neuron can be mathematically described with this function:

$$f(a_1, \dots, a_n) = \sigma \left(b + \sum_{i=1}^n a_i w_i \right) \quad (1)$$

where σ is an activation function which can be either smooth sigmoid function like logistic function

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

¹The research project was supported by Grant No. IGA F4/54/2015 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

, hyperbolic tangens

$$\sigma(x) = \tanh(x) \quad (3)$$

, or stepwise function like

$$\sigma(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (4)$$

3 DATA EXTENSION PHASE

The data consists of several time series imported using R's *quantmod* package [1]. One of the uses of this package is a Yahoo Finance API to import daily stock market data. At first, simple logarithmization and differentiation of time series is used. X_t is defined as the price of the stock

$$x_t := \log X_t \quad (5)$$

$$\Delta x_t = x_t - x_{t-1} \quad (6)$$

We define two windows and their sizes, w and k . w is the length of a window from which we extend the predictor variables and which we call a *lookbehind*. k denotes the length of a window which we call a *lookahead*. It is the window which will be used to extract dependent variable - that is the BSI index. n is the number of observations.

3.1 Preprocessing explanatory variables

For each lookbehind window of fixed length, we compute following statistics: mean (7), median (8), mid (9), logarithm of standard deviation (10), logarithm of mean absolute deviation (11), logarithm of median absolute deviation (12). The index of this new series is defined as s .

$$x_s^{mean} = \frac{\sum_{t=s}^{s+w-1} \Delta x_t}{w} \text{ for each } s = 1 \dots n - w - k - 2 \quad (7)$$

$$x_s^{median} = \text{median}_{t=s}^{s+w-1} \Delta x_t \text{ for each } s = 1 \dots n - w - k - 2 \quad (8)$$

$$x_s^{mid} = \frac{\max_{t=s}^{s+w-1} \Delta x_t + \min_{t=s}^{s+w-1} \Delta x_t}{2} \text{ for each } s = 1 \dots n - w - k - 2 \quad (9)$$

$$x_s^{sd} = \ln \left(\sqrt{\frac{1}{w} \sum_{t=s}^{s+w-1} (\Delta x_t - x_s^{mean})^2} \right) \text{ for each } s = 1 \dots n - w - k - 2 \quad (10)$$

$$x_s^{mad} = \ln \left(\frac{1}{w} \sum_{t=s}^{s+w-1} |\Delta x_t - x_s^{mean}| \right) \text{ for each } s = 1 \dots n - w - k - 2 \quad (11)$$

$$x_s^{mad*} = \ln \left(\text{median}_{t=s}^{s+w-1} |\Delta x_t - x_s^{mean}| \right) \text{ for each } s = 1 \dots n - w - k - 2 \quad (12)$$

After that another set of variables were calculated. For each lookbehind window s , three autoregression models were estimated - AR(1), AR(2) and AR(3), with c_s and ϕ_s being regressors and $\epsilon_{s,t}$ being the error term – white noise.

$$\Delta x_t = c_s^1 + \phi_s^{1,1} \Delta x_{t-1} + \epsilon_{s,t}^1 \quad (13)$$

$$\Delta x_t = c_s^2 + \phi_s^{2,1} \Delta x_{t-1} + \phi_s^{2,2} \Delta x_{t-2} + \epsilon_{s,t}^2 \quad (14)$$

$$\Delta x_t = c_s^3 + \phi_s^{3,1} \Delta x_{t-1} + \phi_s^{3,2} \Delta x_{t-2} + \phi_s^{3,3} \Delta x_{t-3} + \epsilon_{s,t}^3 \quad (15)$$

These estimated AR parameters were used as another explanatory variables for the main model. One for each AR model and regressor. Six in total.

$$x_s^{ar(a,b)} = \phi_s^{a,b} \text{ for each } a = 1 \dots 3, b = 1 \dots a \quad (16)$$

Finally last set of nine augmented predictors are calculated for each window using residuals $e_{s,t}^i$ from the previous autoregressive models.

$$x_s^{ar(i)sd} = \ln \left(\sqrt{\frac{1}{w} \sum_{t=s}^{s+w-1} (e_{s,t}^1)^2} \right) \text{ for each } s = 1 \dots n - w - k - 2 \quad (17)$$

$$x_s^{ar(i)mad} = \ln \left(\frac{1}{w} \sum_{t=s}^{s+w-1} e_{s,t}^2 \right) \text{ for each } s = 1 \dots n - w - k - 2 \quad (18)$$

$$x_s^{ar(i)mad} = \ln (\text{median}_{t=s}^{s+w-1} e_{s,t}^3) \text{ for each } s = 1 \dots n - w - k - 2 \quad (19)$$

3.2 Preprocessing dependent variable

BSI index is estimated with equation (21). We look ahead up to a fixed amount of time (lookahead window) and calculating if there is a time in which selling a stock would yield a profit. Model takes into account transaction prices (q for buy and p for sell) and a threshold of determination Θ which determines minimum amount of yield. If the model predicts BSI as 1, it recommends buying the stock. The X_i is defined from the first section as the price of the stock.

$$\left(\frac{X_t (1+q)}{X_{t+i} (1-p)} \right)^{\frac{365}{i}} \geq \Theta \quad (20)$$

$$BSI_s = \begin{cases} 0 & \max_{j=1+w}^{k+w} \left(\frac{X_t (1+q)}{X_{t+j} (1-p)} \right)^{\frac{365}{j}} \leq \Theta \\ 1 & \max_{j=1+w}^{k+w} \left(\frac{X_t (1+q)}{X_{t+j} (1-p)} \right)^{\frac{365}{j}} \geq \Theta \end{cases} \text{ for each } s = 1 \dots n - w - k - 2 \quad (21)$$

4 ANN MODEL ARCHITECTURE

The BSI dependant variable was predicted with a neural network, which had 21 neurins in a input layer, three hidden layers with 40, 30 and 10 neurons and a single-neuron output layer. Cross entropy error function was used, which has the following form:

$$H(p, q) = - \sum_i p_i \log q_i = -y \log \hat{y} - (1 - y) \log(1 - \hat{y}) \quad (22)$$

$$L(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N H(p_n, q_n) = -\frac{1}{N} \sum_{n=1}^N \left[y_n \log \hat{y}_n + (1 - y_n) \log(1 - \hat{y}_n) \right], \quad (23)$$

where H is a measure of simmilarity between p and q , and $L(\mathbf{w})$ is a cross entropy loss function. A logistic function was used as neuron activation function.

5 CROSS-VALIDATION AND RESULTS

The estimation uses Cross-Validation as another method to deal with over fitting.

The model was fitted and cross-validated on data of several Nasdaq stock market series, namely VIAB, GOOG, FB and NFLX. The area under the curve (AUC) of ROC of the logistic binary classifier was used as a model quality measure. Model parameters were estimated on the sample of 1000 data points from those series. Number of cross-validation sets were 5. After that, out of sample test on the SBUX stock was predicted and its AUC was computed.

Figure 1 shows example ROC curve of prediction using the parameters $w = 7; k = 7; \Theta = 0.2$ on SBUX out of sample dataset. The area under the curve is 0.54.

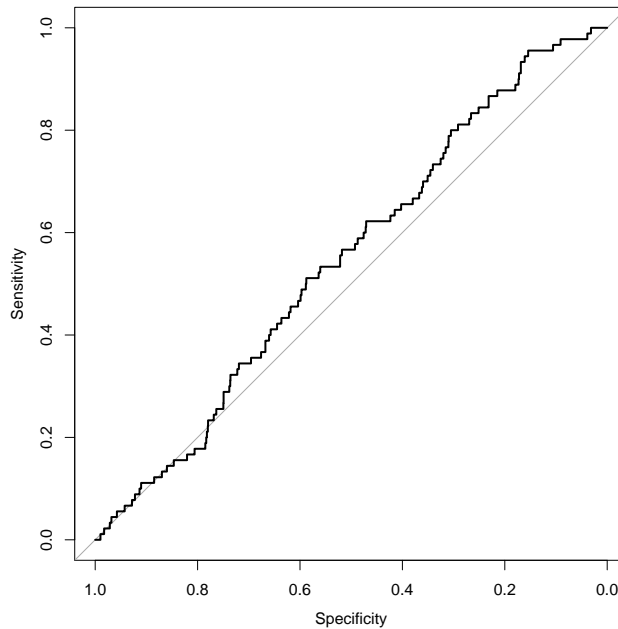


Figure 1: ROC curve of VIAB out of sample test dataset. Source: own R output

6 CONCLUSION

I have introduced a model useful for prediction of univariate time series. This paper has described dependent and explanatory variables as well as method for estimation. Since the dependent was a binary variable, area under the curve of receiver operating characteristic was used as a measure of quality. Further discussions may focus on the threshold of predicted variables that would yield best investment. It is also possible to focus on emulating the actual market behavior of this model and observe its results.

References

- [1] RYAN, J. A. (2016). `quantmod`: Quantitative financial modelling framework. *R package version 0.3-5*. URL <http://www.quantmod.com> URL <http://r-forge.r-project.org/projects/quantmod>.
- [2] WERBOS, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences.

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

ZOBECNĚNÍ BRAESSOVA PARADOXU

GENERALIZATION OF BRAESS PARADOX

Petr Fiala

Abstrakt

V současnosti existuje řada důležitých síťových systémů. Na sítích vznikají paradoxní situace. Mezi klasické paradoxy patří Braessův paradox, kdy přidání dalšího spojení může vést ke zhoršení situace na síti. V příspěvků jsou analyzována zobecnění Braessova paradoxu. Je analyzován vícekriteriální případ s emisemi a dynamická verze Braessova paradoxu. Tato zobecnění vykazují ještě silnější verzi Braessova paradoxu.

Klíčová slova: síťový systém, Braessův paradox, emise, dynamika

Abstract

Currently there are a number of important network systems. Paradoxical situations arise on the networks. Braess paradox belongs among classic paradoxes when adding additional connections may lead to deteriorations in the network. The paper analyzes generalizations of Braess paradox. Multicriterial case with emissions and dynamic version of Braess paradox are analyzed. These generalizations have an even stronger version of Braess paradox.

Keywords: network system, Braess paradox, emissions, dynamics

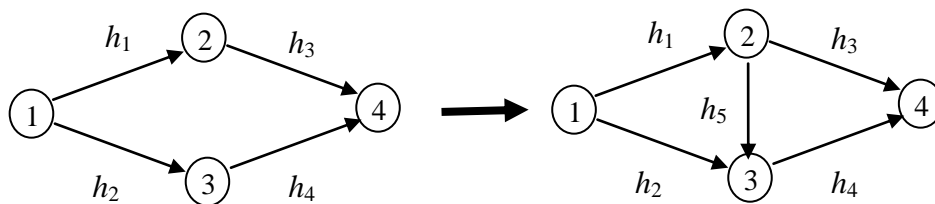
1 ÚVOD

Mnoho současných síťových systémů je charakteristických jak rozsáhlostí, tak i složitostí síťové topologie. Přetížení hraje rostoucí roli nejen v dopravních sítích, ale také v telekomunikačních sítích. Komplementarita se stává rozhodujícím faktorem na informačních trzích. Sítě vykazují pozitivní externalitu. Hodnota jednotky zboží roste s očekávaným počtem prodaných jednotek. Významné jsou také náklady na přepnutí při přechodu na jiné služby nebo přechodu na novou technologii. Rozhodování uživatelů sítě neovlivňuje jen je samotné, ale i ostatní účastníky ve velikosti zisku a nákladů, včasnosti dodávek, kvality životního prostředí atd. Na sítích vznikají paradoxní situace. Mezi klasické paradoxy patří Braessův paradox. V příspěvků jsou analyzována zobecnění Braessova paradoxu. Je analyzován vícekriteriální případ s emisemi a dynamická verze Braessova paradoxu.

2 BRAESSŮV PARADOX

Chování účastníků sítě může být nekooperativní. Příkladem může být chování uživatelů dopravních nebo telekomunikačních sítí, kdy optimalizace z hlediska jednotlivých uživatelů nevede k optimalizaci z hlediska celého systému. Tuto situaci ilustruje známý Braessův paradox, kdy přidání dalšího spojení při stejné poptávce po službě vede při optimalizaci z hlediska uživatelů ke zvýšení nákladů pro všechny (Braess, 1969).

Uvedme konkrétní ilustrační příklad:



Obr. 1. Braessův paradox

Původní síť na Obr. 1 se skládá ze čtyř uzlů 1, 2, 3, 4 a čtyř hran h_1, h_2, h_3, h_4 , začátek a konec sítě je tvořen uzly 1 a 4. Existují dvě cesty mezi začátkem a koncem sítě $C_1 = \{h_1, h_3\}$ a $C_2 = \{h_2, h_4\}$. Předpokládejme, že náklady na jednotlivých hranách v závislosti na velikosti toků jsou

$$n_1(x_1) = 10 x_1, n_2(x_2) = x_2 + 50, n_3(x_3) = x_3 + 50, n_4(x_4) = 10 x_4$$

a celkový požadovaný tok sítě $X = 6$.

V případě optimalizace z pohledu uživatele je rovnovážné řešení dáno situací, kdy všechny cesty mezi začátkem a koncem sítě mají stejné minimální náklady a tudíž uživatelé nemají zájem měnit toky na cestách. Rovnovážné řešení je dáno toky na hranách

$$x_1^* = 3, x_2^* = 3, x_3^* = 3, x_4^* = 3$$

a odpovídajícími náklady na cestách

$$n(C_1) = 83, n(C_2) = 83.$$

Po změně sítě přidáním hrany h_5 mezi uzly 2 a 3 s náklady $n_5(x_5) = x_5 + 10$ vznikne další cesta $C_3 = \{h_1, h_5, h_4\}$ a řešení rovnovážné v původní síti nebude již v této situaci rovnovážné. Rovnovážné řešení bude posílat po každé ze tří cest velikost toku 2, a tudíž na jednotlivých hranách budou toky

$$x_1^* = 4, x_2^* = 2, x_3^* = 2, x_4^* = 4, x_5^* = 2$$

a odpovídajícími náklady na cestách

$$n(C_1) = 92, n(C_2) = 92, n(C_3) = 92.$$

Náklady vzrostly pro každého uživatele sítě z hodnoty 83 na hodnotu 92. Toto zvýšení je také způsobeno skutečností, že hrany h_1 a h_4 jsou sdíleny dvěma cestami a rostou na nich toky a náklady. Braessův paradox je vyvolán danou topologií sítě a rovněž chováním účastníků, které je zaměřeno na optimalizaci z hlediska individuálního uživatele. Pokud je přidána cesta mezi začátkem a koncem sítě, která nesdílí hrany s původním spojením, potom Braessův paradox nevzniká.

3 BRAESSŮV PARADOX S EMISEMI

Je možno uvažovat i vícekriteriální formulaci. Úlohu s náklady je možno rozšířit o uvažování emisí spojených s toky na hranách. Přidání hrany s nulovými emisemi vede ke zvýšení celkových emisí beze změny v poptávce (Nagurney, Dong, 2002).

Uvažujme příklad z původní sítě na Obr. 1, se stejným zadáním. Předpokládejme navíc, že emise na jednotlivých hranách v závislosti na velikosti toků jsou

$$e_1(x_1) = 2 x_1, e_2(x_2) = x_2, e_3(x_3) = x_3, e_4(x_4) = 2 x_4.$$

Pro celkový požadovaný tok sítě $X = 6$ je původní rovnovážné řešení dáno toky na hranách

$$x_1^* = 3, x_2^* = 3, x_3^* = 3, x_4^* = 3$$

a celkové emise jsou

$$E = 2(3) + 1(3) + 1(3) + 2(3) = 18.$$

Po změně sítě přidáním hrany h_5 mezi uzly 2 a 3 s nulovými emisemi bude rovnovážné řešení dáno toky na hranách

$$x_1^* = 4, x_2^* = 2, x_3^* = 2, x_4^* = 4, x_5^* = 2$$

a celkové emise se zvýší

$$E = 2(4) + 1(2) + 1(2) + 2(4) = 20.$$

Přidání nové hrany s nulovými emisemi tedy nejen zvýší náklady pro všechny účastníky, ale i množství generovaných emisí.

4 DYNAMICKÁ VERZE BRAESSOVA PARADOXU

Úlohu je možno zdynamizovat (Nagurney, 2006). Budeme vycházet ze situace, která je popsána na Obr. 1 po změně, která vznikla přidáním hrany h_5 . Nyní však bude poptávka časově závislá

$$X(t) = t, t \in \langle 0, T \rangle$$

a na jednotlivých hranách budou časově závislé toky

$$x_1(t), x_2(t), x_3(t), x_4(t), x_5(t)$$

s náklady, které jsou funkcemi časově závislých toků

$$n_1(x_1(t)) = 10 x_1(t), n_2(x_2(t)) = x_2(t) + 50, n_3(x_3(t)) = x_3(t) + 50, \\ n_4(x_4(t)) = 10 x_4(t) \text{ a } n_5(x_5(t)) = x_5(t) + 10, t \in \langle 0, T \rangle.$$

Vzhledem k topologii sítě a lineárním nákladovým funkcím je tento dynamický případ rovnováhy snadno řešitelný. Obecné případy je možno řešit využitím evolučních variačních nerovnic.

V síti existují 3 cesty $C_1 = \{h_1, h_3\}$, $C_2 = \{h_2, h_4\}$ a $C_3 = \{h_1, h_5, h_4\}$, při velikosti celkového toku t dostáváme 3 rovnovážná řešení:

1. Pouze cesta s tokem o velikosti t . Na jednotlivých hranách jsou toky

$$x_1(t) = t, \quad x_2(t) = 0, \quad x_3(t) = 0, \quad x_4(t) = t, \quad x_5(t) = t$$

s náklady na této cestě

$$n(C_3) = 21t + 10.$$

2. Cesty C_1 a C_2 s tokem o velikosti $t/2$ na každé cestě. Na jednotlivých hranách jsou toky

$$x_1(t) = t/2, \quad x_2(t) = t/2, \quad x_3(t) = t/2, \quad x_4(t) = t/2, \quad x_5(t) = 0$$

s rovnovážnými náklady na každé cestě

$$n(C_1) = n(C_2) = (11/2)t + 50.$$

3. Cesty C_1 a C_2 s tokem o velikosti $(11/13)t - (40/13)$ na každé cestě a cesta C_3 s tokem $(80/13) - (9/13)t$. Na jednotlivých hranách jsou toky

$$\begin{aligned} x_1(t) &= (2/13)t + (40/13), & x_2(t) &= (11/13)t - (40/13), \\ x_3(t) &= (11/13)t - (40/13), & x_4(t) &= (2/13)t + (40/13), \\ x_5(t) &= (80/13) - (9/13)t \end{aligned}$$

s rovnovážnými náklady na každé cestě

$$n(C_1) = n(C_2) = n(C_3) = (31/13)t + (1010/13).$$

Analýzou dynamicky vyjádřených nákladů je možno rozdělit interval pro celkovou poptávku t na tři intervaly:

1. Pro $t \in \left\langle 0, \frac{80}{31} \right\rangle$ je využita pouze cesta C_3 , která využívá přidanou novou hranu h_5 .
2. Pro $t \in \left(\frac{80}{31}, \frac{80}{13} \right)$ jsou využity všechny tři cesty, včetně nové cesty C_3 a vzniká Braessův paradox.
3. Pro $t \in \left(\frac{80}{13}, \infty \right)$ jsou využity pouze staré cesty C_1 a C_2 a přidaná nová hrana h_5 není využita vůbec.

Dynamická verze problému ukázala ještě silnější verzi Braessova paradoxu, pro určité hodnoty poptávky v čase (dynamického celkového toku) Braessův paradox vzniká a pro některé hodnoty poptávky v čase dokonce nebude vůbec využita nová cesta.

5 ZÁVĚR

Braessův paradox ukazuje zajímavou situaci, která vznikne přidáním dalšího spojení v síťovém systému. Analyzovaná zobecnění vykazují ještě silnější verze Braessova paradoxu. Tato zobecnění budou i tématem pro další výzkum.

Výzkumný projekt je podporován grantem č. P402/12/G097 Grantové agentury České republiky a grantem č. IGA F4/54/2015, Fakulty informatiky a statistiky, VŠE, Praha.

Použitá literatura

1. BRAESS, D. (1968). Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung* 12, 258 – 268.
2. FIALA, P. (2016). Dynamické vytváření cen a alokace zdrojů v sítích. Professional Publishing, Praha
3. NAGURNEY, A., DONG, J. (2002). *Supernetworks: Decision-Making for the Information Age*. Edward Elgar Publishers, Cheltenham.
4. NAGURNEY, A. (2006). Supply chain network economics: dynamics of prices, flows and profits. Edward Elgar Publishers, Cheltenham.

Kontaktní údaje

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

Vysoká škola ekonomická, Fakulta informatiky a statistiky
nám. W. Churchilla 4, 130 67 Praha 3

Tel: (420 2) 24 095 447

email: pfiala@vse.cz

POUŽITÍ LOGISTICKÉ REGRESE PRO ODHAD PARAMETRŮ POLOŽEK PŘI ZMĚNÁCH STRUKTURY LATENCE

USING LOGISTIC REGRESSION FOR ITEM PARAMETER ESTIMATION IN THE CASE OF CHANGES IN LATENT TRAIT STRUCTURE

Lenka Fířtová

Abstrakt

Modely teorie odpovědi na položku, které se často používají v high-stake testování, obvykle vyžadují, aby měřený latentní rys byl unidimenzionální. V praxi testy zřídka měří jeden jediný latentní rys, ale spíše množinu několika silně provázaných latentních rysů. Pro úspěšné napojení testů na jednu škálu by mělo platit, že tato množina latentních rysů je napříč propojovanými testy stabilní. Tak tomu však vždy není. Tento článek proto navrhuje metodu použitelnou k propojení testů v případě nestability měřených latentních rysů.

Klíčová slova: teorie odpovědi na položku, unidimenzionalita latence, odlišné fungování položek

Abstract

Item response theory models, which are frequently used in high-stake testing, often require the latent trait being measured to be unidimensional. In reality, tests seldom measure one single trait, but rather a set of several highly intercorrelated traits. In order to successfully link different tests to the same scale, this set of traits should remain stable across the tests being linked. This is not always the case though. Therefore, this paper suggests a method that can be used to link tests in the presence of latent trait structure instability.

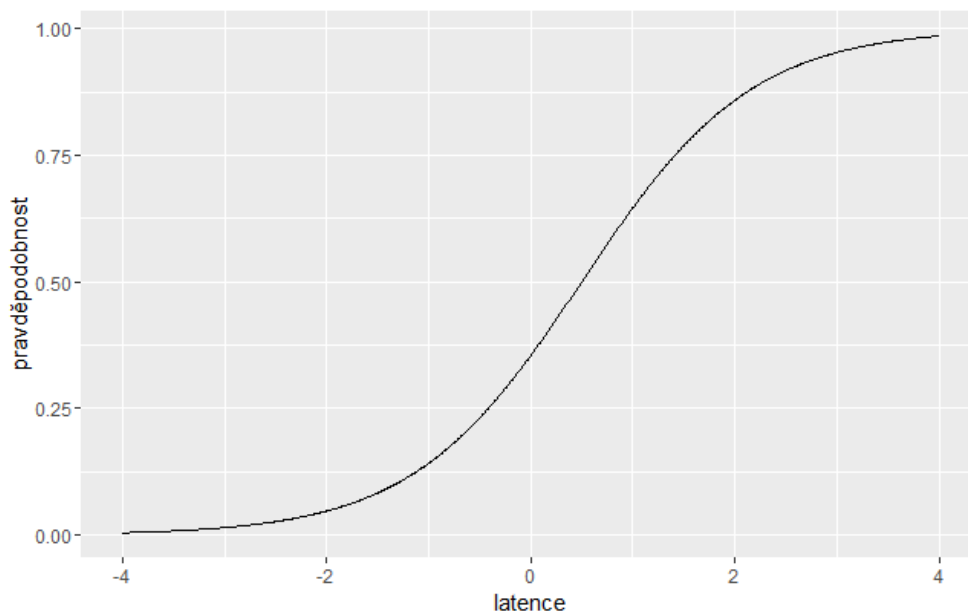
Keywords: item response theory, unidimensionality of the latent trait, differential item functioning

1 TEORIE ODPOVĚDI NA POLOŽKU

Teorie odpovědi na položku (*item response theory*, dále *IRT*) je souborem modelů a postupů, které se používají v kontextu testování (ve vzdělávání, psychologii apod.). V rámci teorie odpovědi položku byla vyvinuta celá řada modelů, avšak jedním z nejběžnějších dichotomických modelů je unidimenzionální dvouparametrický model (Birnbaum, 1968), který modeluje pravděpodobnost, že respondent i s úrovní latentního rysu (latence) θ_i odpoví správně na položku j jako

$$P(x_{ij} = 1 | \theta_i, b_j, a_j) = \frac{1}{1 + e^{-a_j(\theta_i - b_j)}}, \quad (1)$$

Kde $x_{ij} = 1$ resp. 0, zodpoví-li respondent i položku j správně resp. nesprávně, a b_j , a_j značí obtížnost resp. diskriminaci položky j . Latencí se rozumí úroveň měřené dovednosti respondenta (například matematická gramotnost). Vztah mezi úrovní latence a pravděpodobností správně odpovědi zobrazuje tzv. charakteristická křivka položky, příklad znázorňuje graf 1.



Graf 1: Charakteristická křivka položky

Obtížnost položky je v kontextu *IRT* definována jako taková úroveň latence, při níž má respondent právě 50% pravděpodobnost, že položku zodpoví správně, tzn. parametr obtížnosti určuje horizontální posun charakteristické křivky.

Obtížnost položky a latence respondentů jsou tudíž na jedné škále. Diskriminace položky určuje sklon křivky: čím je vyšší, tím lépe položka rozlišuje mezi horšími a lepšími respondenty.

Kromě dichotomických modelů existují i modely polytomické, které se nejčastěji využívají v psychologických dotaznících, kde respondenti odpovídají na Likertově škále. Existují rovněž mutlidimenzionální *IRT* modely, které modelují výše uvedenou pravděpodobnost jako funkci více než jedné latence.

V kontextu *educational measurement* (například u znalostních testů či testů studijních předpokladů) se však typicky vyskytují položky dichotomické (respondent zde obvykle může odpovědět buď správně, nebo nesprávně), a zároveň zde většinou existuje požadavek na unidimenzionalitu latence, protože výsledkem testu má být jedno jediné skóre, proto je výše uvedený model v této oblasti jedním z nejčastěji používaných.

2 POPIS PROBLÉMU

Jedním z předpokladů unidimenzionálních *IRT* modelů je předpoklad unidimenzionality latentního rysu, jinak řečeno předpoklad, že všechny položky testují právě jednu a tutéž latenci. V případě dvouparametrického modelu je tento předpoklad obzvláště důležitý, protože parametry všech položek v testu se odhadují najednou. Ve chvíli, kdy položky netestují stejný latentní rys, se model snaží zachytit nějakou společnou, dominantní latenci, která všechny položky testu propojuje, a právě k této latenci se pak budou odhady všech parametrů vztahovat. V realitě je ovšem předpoklad unidimenzionality latentního rysu splněn jen málokdy.

Je-li naším cílem odhad parametrů položek a osob pouze z jednoho testování, a nikoli vzájemné propojování několika testů, pak mírné porušení předpokladu unidimenzionality latence nepředstavuje zásadní problém. Pokud totiž test obsahuje nějaký výrazný latentní rys, který určitým způsobem všechny položky propojuje, dovede jej model poměrně dobře podchytit. Toto tvrzení podporují závěry studie (Drasgow, Parsons, 1983), v níž bylo zkoumáno použití unidimenzionálního dvouparametrického modelu v pěti typech testů zaměřených na různě silně korelované latentní rysy: od zcela unidimenzionálního latentního rysu, po pět téměř nezávislých latentních rysů. Podle této studie je dvouparametrický model použitelný, je-li korelace latentních rysů rovna alespoň 0,46, nemusí být tedy nutně velmi vysoká.

Situace je složitější v případě, kdy dochází k propojování testů, typicky za účelem převodu parametrů na jednotnou škálu. Propojování testů se používá zejména u high-stake testů typu přijímací zkoušky, kde je potřeba zajistit srovnatelnost výsledků (v rámci roku, meziročně apod.), a provádí se většinou přes společné (tzv. kotvící) položky. Parametry kotvících položek přitom ovšem musí zůstat neměnné napříč testy, přičemž existuje řada postupů, jak tzv. odlišné chování položky (*differential item functioning*, dále *DIF*) otestovat. *DIF* může mít řadu příčin, například odlišnou časovou dotací na test, odlišné umístění položky v testu apod. Ve chvíli, kdy každý z testů testuje jinou směs latence, lze rovněž často vysledovat *DIF*, neboť parametry položek v testu se vždy budou vztahovat k příslušné směsi latence, na niž je daný test zaměřen, a pokud se tato směs latence napříč testy liší, pak lze očekávat, že se budou lišit i parametry kotvících položek.

Je tedy zřejmé, že odhad parametrů všech položek testu zároveň, tak jako se to u dvouparametrického modelu obvykle provádí, může být ve chvíli, kdy se směs latence napříč testy mění, problematický, protože parametry položek, které jsou společné více testům, budou při změnách měřené směsi latence pravděpodobně vykazovat nekonzistentní chování napříč testy. Chceme-li v tomto případě zachovat latenci, k níž se položky společné těmto testům vztahují, je potřeba postupovat jiným způsobem.

3 NÁVRH ŘEŠENÍ

Uvažujme situaci, kdy máme test o 40 položkách. Prvních 20 položek pochází ze staršího testování a měří latenci, kterou označme I. (tzn. tyto položky byly obsaženy v některém starším testu, který byl celý zaměřen na latenci I.). Druhých 20 položek jsou položky nového typu, které částečně měří původní latenci I., ale zároveň je k jejich řešení potřeba i nějaká další dovednost, latence II. Příkladem může být test z angličtiny: pokud původní test měřil gramatiku a nové položky jsou rovněž zaměřeny na gramatiku, ale v kontextu delšího textu, měří nové položky kromě znalostí gramatiky i schopnost porozumění delšímu textu.

Prvním možným, a typicky používaným postupem (dále nazýván metoda A) je odhadnout všechny položky zároveň. Tím získáme odhady parametrů položek, které se vztahují ke směsi latencí I. a II. Parametry prvních 20 položek v testu, které měří pouze latenci I., budou mít pravděpodobně nižší odhad diskriminace ve srovnání se situací, kdy by celý test měřil pouze latenci I., protože latenci II., která je v testu jako celku také měřena, neměří. Pokud nám jde pouze o odhad parametrů tohoto jednoho testu, jedná se v případě relativně silné korelace latencí I. a II. o přijatelný postup. Ovšem při propojování několika testů může být tento postup problematický, neboť položky společné několika testům mohou vykazovat nekonzistentní chování napříč testy z důvodu změny směsi latence, k níž se odhad jejich parametrů vztahuje.

Druhý možný postup (dále nazýván metoda B) je následující: odhadneme parametry pouze prvních 20 položek v testu (případně je rovnou převezmeme ze staršího testování, to závisí na

metodě propojování testů, tzv. harmonizace). Poté odhadneme latenci osob pouze za použití těchto 20 položek. Získáme tak odhad jejich úrovně latence I., bez ohledu na to, jaká je jejich úroveň latence II. Tento odhad latence I. pak použijeme jako vysvětlující proměnnou pro odhad parametrů zbylých 20 položek z testu, jejichž parametry odhadneme logistickou regresí. Takto odhadnuté parametry položek se tedy budou vztahovat k latenci I., ačkoli měří ještě další latenci. Ve svém důsledku to znamená, že odhad parametrů všech položek v testu bude vztažen k latenci I., tedy k latenci, která byla předmětem původního testu, z něhož prvních 20 položek pochází.

3.1 Podrobný popis postupu při použití metody A

V případě metody A, kdy odhadujeme parametry všech 40 položek zároveň, je řešení přímočaré. Odhad parametrů položek a osob se zde provádí metodou maximální věrohodnosti. Ve chvíli, kdy nejsou známy parametry položek ani osob, probíhá odhad simultánně, a to tak, že se určí prvotní odhad parametrů položek (například se může vyjít z prostého podílu osob, který položku zodpověděl správně), metodou maximální věrohodnosti se odhadne latence osob při takto zafixovaných parametrech položek, na základě takto odhadnuté latence se metodou maximální věrohodnosti odhadnou parametry položek, a tak dále, dokud není dosaženo konvergence, přičemž funkce v R vrací při tomto postupu pouze odhady parametrů položek (odhad latence osob se zde provádí zvlášť jako následný krok).

Jedním ze základních problémů *IRT* je neurčenost metriky: obvykle se postupuje tak, že se zafixuje střední hodnota latence osob na 0 a rozptyl na 1 (s tímto předpokladem pracují i balíčky pro R).

Označme \hat{b}_j resp. \hat{a}_j takto získaný odhad parametru obtížnosti resp. diskriminace j -té položky.

Dále definujme $\hat{P}_{ij} \equiv P(x_{ij} = 1 | \theta_i, \hat{b}_j, \hat{a}_j)$. Poté, co jsme odhadli parametry položek, můžeme odhadnout latenci osob nalezením argumentu maxima následující věrohodnostní funkce (většinou se používá její logaritmus):

$$\hat{\theta}_i = \underset{\theta_i}{\operatorname{argmax}} \mathcal{L}(x_{i1} \dots x_{i40} | \theta_i) = \prod_{j=1}^{40} (\hat{P}_{ij})^{x_{ij}} (1 - \hat{P}_{ij})^{1-x_{ij}}. \quad (2)$$

3.2 Podrobný popis postupu při použití metody B

V případě metody B postupujeme v několika krocích. V prvním kroku odhadneme metodou maximální věrohodnosti parametry prvních 20 položek v testu, tedy postupujeme stejně jako u metody A s tím rozdílem, že použijeme pouze prvních 20 položek. Následně odhadneme pro každého z respondentů úroveň jeho latentního rysu jako

$$\hat{\theta}_i = \underset{\theta_i}{\operatorname{argmax}} \mathcal{L}(x_{i1} \dots x_{i20} | \theta_i) = \prod_{j=1}^{20} (\hat{P}_{ij})^{x_{ij}} (1 - \hat{P}_{ij})^{1-x_{ij}}, \quad (3)$$

tedy opět postupujeme stejně jako u metody A. Zde je na místě poznamenat, že pro odhad latence osob existuje několik používaných modifikací výše uvedené věrohodnostní funkce. Jednou z nevýhod „klasické“ metody maximální věrohodnosti je totiž například skutečnost, že pomocí ní nelze odhadnout latenci pro respondenty, kteří na všechny položky odpověděli správně, resp. na všechny nesprávně. Modifikacemi metody maximální věrohodnosti, jsou

například bayesovské metody odhadu (*expected a posteriori*, *maximum a posteriori*), a dále pak tzv. *Warm's weighted likelihood estimate* (dále *WLE*), u něhož Warm (1989) ukázal, že oproti klasické metodě maximální věrohodnosti snižuje vychýlení. V následné simulaci budeme při odhadu latence osob vždy pracovat právě s *WLE*.

Po odhadu latence respondentů na základě prvních 20 položek v testu odhadneme parametry (obtížnost a diskriminaci) zbylých 20 položek v testu jedna položka po druhé logistickou regresí, kde vysvětlující proměnnou bude odhad latence respondentů z prvního kroku postupu.

Označme nyní $P_{ij} \equiv P(x_{ij} = 1 | \theta_i, b_j, a_j)$. Při znalosti skutečné latence respondentů bychom odhadovali regresi ve tvaru:

$$\ln\left(\frac{P_{ij}}{1-P_{ij}}\right) = -a_j b_j + a_j \theta_i, \quad j = 21 \dots 40. \quad (4)$$

To znamená, že hledaným odhadem diskriminace by byl přímo odhad sklonu této logistické křivky a hledaný odhad obtížnosti by pak bylo možné dopočítat z odhadu úrovně konstanty.

Protože skutečnou hodnotu latence neznáme, používáme místo θ_i pouze její odhad $\hat{\theta}_i$, což ovšem znamená, že součástí vysvětlující proměnné je i náhodná chyba, která vznikne při odhadu latence, a odhad parametrů obtížnosti a diskriminace tak bude vychýlený. Proto je potřeba provést korekci, a to tak, že po odhadu regrese (4) vydělíme odhad diskriminace hodnotou

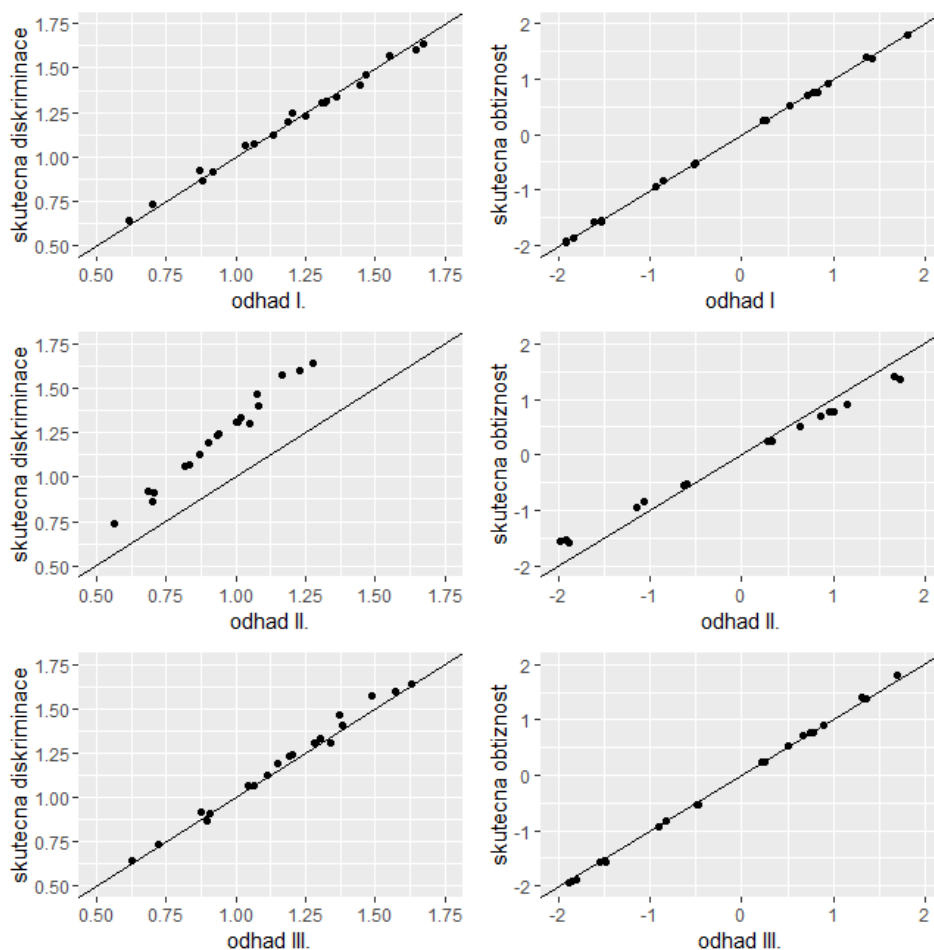
$$\frac{\frac{1}{n} \sum_{i=1}^n [\text{var}(\hat{\theta}_i) - \bar{SE}^2 \hat{\theta}_i]}{\text{var}(\hat{\theta}_i)}, \quad (5)$$

kde n je počet respondentů, načež s využitím takto korigovaného odhadu diskriminace dopočítáme odhad obtížnosti. Uvedená hodnota vychází z toho, že chybu ve vysvětlujících proměnných je možné ošetřit vydělením odhadnutých parametrů reliabilitou. V tzv. klasické teorii testů se reliabilita vyjadřuje jako podíl rozptylu skutečného skóre ($\text{var}T$) a rozptylu napozorovaného skóre ($\text{var}X$), kde rozptyl napozorovaného skóre je součtem rozptylu skutečného skóre a rozptylu náhodné chyby ($\text{var}X = \text{var}T + \text{var}E$). Tak lze získat hodnotu reliability (ρ) pro celou skupinu respondentů, tzn.

$$\rho = \frac{\text{var}T}{\text{var}X} = \frac{\text{var}X - \text{var}E}{\text{var}X}. \quad (6)$$

Potřebné hodnoty většinou neznáme, a proto reliabilitu pouze odhadujeme, například pomocí Cronbachova alfa, nebo metody *split-half*. V kontextu *IRT* lze odhadnout reliabilitu pro každého respondenta zvlášť. Rozptyl napozorovaného skóre ($\text{var}X$) můžeme odhadnout jako rozptyl odhadu latence. Rozptyl náhodné chyby ($\text{var}E$) můžeme odhadnout jako druhou mocninu odhadu směrodatné chyby odhadu latence. Ta se pro jednotlivé respondenty liší, protože latenci osob, pro něž je test obtížnostně přiměřený, lze odhadnout přesněji než latenci velmi nadprůměrných či velmi podprůměrných respondentů. Je tedy zřejmé, že dostaneme pro každého respondenta jinou hodnotu reliability. Pro korekci lze pak použít její průměrnou hodnotu.

Nutnost korekce ilustrujeme na simulaci. Budou zde použity funkce obsaženy v balíčcích *ltm* (Rizopoulos, 2006) a *irtosys* (Partchev, 2008). Vygenerujeme 40 položek s obtížností z rovnoměrného rozdělení $U(-2; 2)$ a diskriminací z normálního rozdělení $N(1,25; 0,25)$. Parametry volíme tak, aby odpovídaly realitě high-stake testů. Poté vygenerujeme latenci pro 10 000 respondentů z normálního rozdělení $N(0; 1)$ a matici odpovědí na uvedené položky. Všechny položky tedy měří jednu a tutéž latenci. Latenci respondentů odhadneme pouze na základě prvních 20 položek v testu. Parametry zbylých 20 položek v testu odhadneme logistickou regresí třemi různými způsoby: zaprvé se skutečnou latencí jakožto vysvětlující proměnnou, zadruhé s odhadem latence jakožto vysvětlující proměnnou (bez další korekce) a zatřetí s odhadem latence jakožto vysvětlující proměnnou s následnou korekcí popsanou výše (dále nazváno odhad I., odhad II., odhad III.). Provedeme 50 replikací a v každé z nich spočítáme pro 20 položek, jejichž parametry odhadujeme logistickou regresí, *RMSE* pro obtížnost a diskriminaci. Výsledek jedné z replikací znázorňuje graf níže, průměrnou hodnotu *RMSE* a její směrodatnou chybu pak tabulka níže. Z grafu i tabulky je patrné, že nejlepších výsledků dosáhneme, jak bychom nejspíš očekávali, při použití skutečných hodnot latence. Při použití odhadů latence bez další korekce jsou parametry diskriminace i obtížnosti systematicky vychýlené. Korekce však vychýlení významně snižuje. Dále budeme proto při používání logistické regrese pracovat s výše popsanou korekcí.



Graf 2: Výsledky jedné z replikací při použití logistické regrese s korekcí a bez korekce

	odhad I.	odhad II.	odhad III.
<i>RMSE obtížnosti</i>	0,03 (0,01)	0,25 (0,04)	0,06 (0,02)
<i>RMSE diskriminace</i>	0,03 (0,01)	0,28 (0,03)	0,05 (0,01)

Tabulka 1: Srovnání RMSE při použití logistické regrese s korekcí a bez korekce

4 POPIS SIMULACE

Nyní provedeme simulaci, v níž porovnáme výsledky metody *A* a výsledky metody *B* (s korekcí). Vygenerujeme 40 položek s obtížností z rovnoměrného rozdělení $U(-2; 2)$ a diskriminací z normálního rozdělení $N(1,25; 0,25)$. Budeme uvažovat dvě situace: v první situaci měří všech 40 položek v testu jednu a tutéž latenci (latence I.), která pochází z $N(0; 1)$. Ve druhém případě měří prvních 20 položek pouze latenci I. z rozdělení $N(0; 1)$, zatímco druhých 20 položek měří částečně latenci I., ale částečně i jinou, související latenci II., která

pochází z normálního rozdělení se střední hodnotou 0 a rozptylem $\left(\frac{-0,5 + \sqrt{12 \cdot 25}}{2}\right)^2$, kde korelace latencí I. a II. je 0,25. Výsledná latence, kterou těchto 20 položek měří, je váženým součtem obou latencí s vahami 0,5. V praxi to znamená, že tato latence měří stejnou vahou původní latenci, ale zároveň i latenci jinou, nýbrž související, což odpovídá realitě testování, viz příklad s testem z anglického jazyka výše: lze očekávat, že znalost gramatiky a porozumění textu spolu budou korelovat a test může obsahovat položky, které měří obojí. Rozptyl latence II. byl zvolen tak, aby (vzhledem ke zvolené korelaci a vahám) výsledná latence měla rozptyl 1, tzn. pocházela z $N(0; 1)$, a to proto, že *R* s tímto předpokladem pracuje, a není-li splněn, musí se obcházet transformacemi.

Vygenerujeme takto popsané latence pro 10 000 respondentů, načež vygenerujeme dvě matice odpovědí, jednu pro první situaci, druhou pro druhou situaci. V obou případech použijeme pro odhad latence položek dvě výše zmiňované metody, *A* a *B*. Při použití metody *A* tedy odhadneme parametry všech 40 položek zároveň a poté odhadneme latenci respondentů metodou *WLE*. Při použití metody *B* odhadneme nejprve parametry pouze prvních 20 položek, z nich odhadneme latenci respondentů metodou *WLE*, a tu poté použijeme jako vysvětlující proměnnou pro odhad parametrů zbylých 20 položek logistickou regresí, načež teprve odhadneme metodou *WLE* výslednou latenci respondentů na základě všech 40 položek. Provedeme 50 replikací.

5 VÝSLEDKY

V tabulce 2 jsou prezentovány následující výsledky: korelace latence respondentů odhadnuté na základě celého testu s jejich skutečnou hodnotou latence I. (dále „korelace“), *RMSE* obtížnosti a diskriminace prvních dvaceti položek v testu (dále „*RMSE obtížnost*“ a „*RMSE diskriminace*“), střední absolutní odchylka a střední odchylka skutečných parametrů zbylých

20 položek v testu a jejich odhadů (dále „*abs. odchylka obtížnost*“, „*abs. odchylka diskriminace*“, „*odchylka obtížnost*“, „*odchylka diskriminace*“). Záměrně zde není použito slovo „chyba“, nýbrž „odchylka“, protože ve druhé situaci se skutečné parametry těchto položek vztahují ke směsi latencí I. a II., zatímco jejich odhady jsou navázány na latenci I., a proto nejsou vzájemně přímo srovnatelné. Střední odchylka nikoli v absolutní hodnotě je uvedena proto, že zejména u parametru diskriminace lze očekávat systematické vychýlení jedním směrem.

	Situace 1		Situace 2	
	Metoda A	Metoda B	Metoda A	Metoda B
<i>korelace</i>	0,95	0,95	0,86	0,89
<i>RMSE obtížnost</i>	0,03	0,04	0,11	0,03
<i>RMSE diskriminace</i>	0,04	0,04	0,17	0,04
<i>abs. odchylka obtížnost</i>	0,03	0,05	0,09	0,37
<i>odchylka obtížnost</i>	0,00	0,00	0,00	0,01
<i>abs. odchylka diskriminace</i>	0,03	0,04	0,16	0,49
<i>odchylka diskriminace</i>	0,00	0,01	0,16	0,49

Tabulka 2 Výsledky simulace

Vidíme, že v první situaci jsou výsledky metody *A* a *B* srovnatelné, ačkoli v některých ohledech jsou výsledky metody *B* nepatrně horší, což je pravděpodobně dáno faktem, že metoda sestává z více kroků, takže zde existuje i více zdrojů chyby.

Ve druhé situaci pozorujeme u metody *A* pokles korelace odhadnuté latence se skutečnými hodnotami latence I., což je logické, neboť test měří navíc i jinou, byť související latenci. Vidíme, že *RMSE* prvních 20 položek se zde významně zvýšila oproti situaci 1, což je zapříčiněno faktem, že i když prvních 20 položek měří stále pouze latenci I., parametry všech 40 položek byly odhadnuty zároveň, a proto se do parametrů prvních 20 položek promítl i fakt, že test jako celek zároveň měří latenci II., byť tyto položky samotné ji neměří. Jinak řečeno, parametry všech položek v testu se nyní vztahují ke směsi latencí I. a II., i přesto, že část položek latenci II. vůbec neměří. Jak již bylo řečeno výše, toto může představovat problém ve chvíli, kde je cílem propojit několik testů a je potřeba zachovat stabilitu parametrů (resp. měřené latence). Co se týče druhých 20 položek, zde pozorujeme rovněž nárůst odchylky oproti situaci 1. To je dáno tím, že test jako celek měří latenci I. ve větší míře, než v jaké ji měří těchto 20 položek, což se opět promítá do odhadu parametrů. Vidíme, že odhad diskriminace je systematicky vychýlen jedním směrem, a to pravděpodobně směrem k nule (odhad diskriminace bude nižší než její skutečná hodnota).

U metody *B* pozorujeme ve druhé situaci mírný nárůst korelace odhadnuté latence se skutečnými hodnotami latence I. oproti metodě *A*, což je dáno tím, že nyní jsou parametry všech položek testu navázány na latenci I. Vidíme dále, že u prvních 20 položek testu je *RMSE* srovnatelná se situací 1, neboť tentokrát není odhad jejich parametrů zanesen latencí II., kterou tyto položky neměří. U posledních 20 položek v testu se naopak odchylka zvětšila, a to proto, že tentokrát jsou odhady jejich parametrů striktně navázány na latenci I., bez ohledu na to, že tyto položky měří zároveň i latenci II.

Výsledky lze tedy shrnout následovně: jestliže všechny položky měří tutéž latenci (situace 1), dávají metody *A* a *B* velmi podobné výsledky, takže lze použít kteroukoli z nich. Ovšem metoda *B* je delší a náročnější, a protože pro její použití v tomto případě není důvod, je vhodnější metoda *A*, tedy odhad parametrů všech položek v testu najednou.

Jestliže položky neměří tutéž latenci (situace 2), je důležité, zda se nám jedná pouze o odhad parametrů položek a osob v daném testu, nebo zda máme zájem testy propojovat (navázat na jednotnou škálu), což je kontextu high-stake testování obvyklejší. V prvním případě lze použít metodu *A*, s tím, že parametry všech položek v testu budou odhadnuty zároveň, a proto se budou všechny vztahovat ke směsi latencí, kterou měří daný test jako celek bez ohledu na to, co měří jednotlivé položky, a i odhad latence osob bude odhadem právě této směsi latencí. Ve druhém případě je ovšem použití metody *A* problematické, protože při propojování testů je nutné, aby se položky vyskytující se ve vícero testech vztahovaly vždy ke stejné směsi latence. Zde je proto vhodnější použít metodu *B*, tedy navázat parametry položek, které měří novou, ve starších testech neobsaženou latenci, právě na tuto latenci za účelem dosažení stability měřené latence napříč testy.

Použitá literatura

1. Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. *Statistical theories of mental test scores*.
2. Drasgow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7(2), 189-199.
3. Partchev, I., Partchev, M. I., & Suggests, M. A. S. S. (2008). The irtoys package.
4. Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of statistical software*, 17(5), 1-25.
5. Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427-450.

Kontaktné údaje

Ing. Lenka Fiřtová

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky

Nám. Winstona Churchilla 4, 130 67, Praha 3

Tel: +420 739 840 803

email: xfirl02@vse.cz

NELINEARITA V DATECH: JE GAM MODEL VŽDY KRÁLEM?

DATA NONLINEARITY: IS ALWAYS GAM KING?

*Lukáš Frýd*¹

Abstrakt

Problematika nelinearity v datech se často řeší pomocí složitých modelů, které často mají neparametrický charakter. V této práci je porovnána predikční schopnost klasického lineárního modelu s GAM modelem na mikroekonomických datech. Test dokazuje, že v případě existence mírné nelinearity, nelze odlišit predikční schopnosti obou modelů.

Klíčová slova: splines, GAM, Diebold-Mariano test

Abstract

Nowadays the data nonlinearity is very popular theme which is solved with complex models. We focus on the prediction ability of GAM and classical linear regression model in microeconomics data. We found that GAM model does not outperform classical linear regression model in case of low nonlinearity.

Keywords: splines, GAM, Diebold-Mariano test

1 ÚVOD

V ekonomie jsou jak díky snadnější interpretaci, tak i nižší výpočetní složitosti, preferovány lineární modely. V reálním světě jsou však vztahy komplexnější a velmi často mají nelineární charakter. Určitým kompromisem může být Generalized Additive Model (GAM), který byl vyvinut autory Hastie a Tibshirani (1986). Tento typ modelu se stal oblíbeným zejména v přírodních vědách pro schopnost postihnout složité nelineární vztahy mezi proměnnými např. Bigelow (1999), He a Laurs (1998). Je proto otázkou, zdali je model GAM vhodnou alternativou i v případech, kdy vztahy mezi proměnnými je popsán “hezkou” funkcí. Z tohoto důvodu je tato práce zaměřena na porovnání GAM modelu s klasickým lineárním modelem na mikroekonomických datech sloužících pro výuku ekonometrie.

2 METODOLOGIE

2.1 Generalized Additive Models

Jedná se o rozšířený lineární regresní model, kde jsou jednotlivé regresory nahrazeny vhodnou nelineární funkcí $f(x)$:

$$y_i = \beta_0 + f_1(x_{i1}) + f_2(x_{i2}) + \dots + f_p(x_{ip}) + \epsilon_i \quad 1)$$

Možnými kandidáty na funkce $f_j(x_{ij})$ mohou být například modely Lokální regrese, spline funkce, skoková funkce, ale je i možnost použít pro popis regresoru například waveletovou

¹ Článek vznikl s podporou projektu IGA F4/73/2016 Interní grantové agentury Vysoké školy ekonomické v Praze.

funkci. Právě flexibilita GAM modelu je jeho hlavní předností. V této práci bude jako vyrovnávací funkce použita metoda vyhlazujících spline funkcí (smoothing Splines).

Princip metody smoothing splines je nejlépe vidět na následujícím optimalizačním algoritmu:

$$\sum_{i=1}^n (y_i - g(x_i))^2 + \lambda \int g''(t) dt \quad 2)$$

Kde λ představuje tzv. tuning parametr a slouží k určení jak moc chceme mít vyhlazenou spline funkci. První část z rovnice 2 je pak klasická ztrátová funkce a druhá část představuje penalizaci. Výběr parametru λ souvisí se stupni volnosti, kdy v této práci bude použita metoda „leave-one-out cross validation error“ (LOOCV).

2.2 Vyhodnocení predikčních schopností modelu

Pro vyhodnocení predikčních schopností obou modelů bude využita metoda podle Diebold a Mariano (1995).

$$DM = \frac{\bar{d}}{\sqrt{V(\bar{d})}} \quad 3)$$

Kde \bar{d} představuje aritmetický průměr z $\mathbf{d}_t = (\mathbf{y}_{t+h}^A - \widehat{\mathbf{y}}_{t+h}^A)^2 - (\mathbf{y}_{t+h}^B - \widehat{\mathbf{y}}_{t+h}^B)^2$, \mathbf{y}_{t+h} je hodnota z out-of-sample vzorku a $\widehat{\mathbf{y}}_{t+h}$ je vyrovnaná hodnota vypočtená z modelu odhadnutém na in-the-sample vzorku. Indexy A a B pak značí rozdílné metody odhadu. Dále $V(\bar{d})$ je robustní odhad pomocí metody Newey – West. DM statistika se v distribuci blíží k normovanému normálnímu rozdělení a $H_0: E(\mathbf{d}_t) = \mathbf{0}$ proti $H_1^A: E(\mathbf{d}_t) < \mathbf{0}$ a $H_1^B: E(\mathbf{d}_t) > \mathbf{0}$.

3 DATA

V práci je použit datový soubor z knihy Introductory Econometrics: A Modern Approach od Woolridge. Jedná se o soubor BWGHT. Datový soubor obsahuje 14 proměnných, kdy je cílem určit statisticky významné faktory ovlivňující váhu novorozence v uncích. Populační regresní funkce je v knize sestavena následovně:

$$bwght = \beta_0 + \beta_1 cigs + \beta_2 \log(faminc) + \beta_3 parity + \beta_4 male + \beta_5 white + \beta_6 mothereduc + \beta_7 fathereduc + \epsilon \quad 4)$$

- Bwght – váha novorozence v uncích
- Cigs – počet cigaret vykouřených za den v období těhotenství
- Faminc – příjem rodiny v roce 1988 v 1000USD
- Parity – pořadí narozeného dítěte v rodině
- Male – pokud je narozené dítě chlapec male=1
- White - pokud je narozené dítě bílé pleti white=1
- Mothereduc – počet let vzdělání matky
- Fathereduc – počet let vzdělání otce

O náhodné složce je zaveden předpoklad, že $\epsilon \sim i. i. d(\mathbf{0}, \sigma^2)$.

4 EMPIRICKÁ ČÁST

4.1 In-the-sample

V empirické části nejprve rozdělíme datový soubor na in-the-sample a out-of-sample. Na vzorku in-the-sample, který bude tvořen 75% dat, odhadneme rovnici 4, dále pak odhadneme následující rovnici pomocí GAM.

$$bweight = \beta_0 + f(cigs) + f(faminc) + f(parity) + male + white + f(mothereduc) + f(fathereduc) + \epsilon \quad 5)$$

Kde $f(x)$ představuje smoothing spline funkci s vhodně zvolenými stupni volnosti, na základě kapitoly 2.1.

Výsledky odhadu pro rovnici 4 jsou v tabulce 1.

Proměnná	Koeficient	Směr.Chyba	t-test	P-value
Intercept	104.09	4.88	21.31	0.00
Cigs	-0.54	0.12	-4.30	0.00
Lfaminc	1.45	1.08	1.33	0.18
Parity	1.91	0.81	2.36	0.02
Male	3.12	1.34	2.34	0.02
White	5.56	1.84	3.01	0.00
Mothereduc	-0.35	0.37	-0.94	0.34
fathereduc	0.47	0.34	1.39	0.16

Tabulka 1. Odhad rovnice 4.

Whittův test ani BP test nesvědčí pro zamítnutí nulové hypotézy o homoskedasticitě. Při otestování sdružené významnosti pro parametry $\beta_2, \beta_6, \beta_7$ nemůžeme zamítnout nulovou hypotézu a proto použijeme omezený model pro hodnocení jeho predikčních schopností v části 4.2.

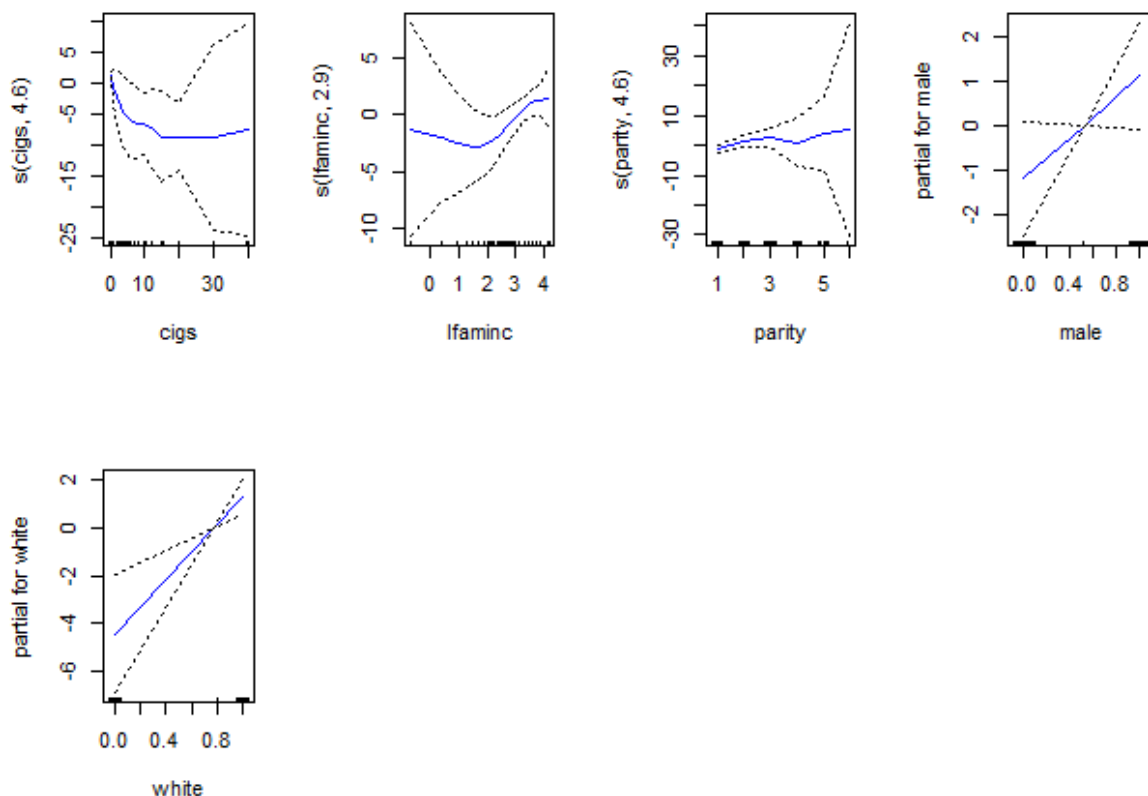
V případě odhadu rovnice 5 pomocí metody GAM vychází, že vzdělání rodičů nemá signifikantní vliv na váhu dítěte. Z tohoto důvodu byl model zredukován do následující podoby:

$$bweight = \beta_0 + s(cigs, 4.6) + s(faminc, 2.9) + s(parity, 4.6) + male + white + \epsilon \quad 6)$$

Kde $s()$ představuje funkci smoothing splines a číslo v závorce odhadnutý počet stupňů volnosti pomocí LOOCV. V rámci tohoto odhadu se jeví, že všechny proměnné mají statisticky významný vliv na váhu novorozence pro $\alpha = 0.01$, pro proměnnou *cigs*, *lfaminc* a *white*. Dále můžeme zamítnout nulovou hypotézu o nevýznamnosti parametrů spojených s *parity* a *male* pro $\alpha = 0.1$. V tomto případě se jedná o testy lineárních vztahů, v případě nelineárních vztahů, nemůžeme říci ani o jedné z proměnných aproximovaných pomocí smooth splines, že by byla statisticky významná. Při pohledu na obrázek 1, charakterizující závislost váhy dítěte na vývoji regresoru můžeme vidět, že proměnná *cigs* a *lfaminc* vykazují

nelineárních charakter. Naopak u proměnné parity se jedná spíše o lineární vztah. Z tohoto důvodu bude pro hodnocení predikčních schopností GAM modelu použita následující verze modelu:

$$bwght = \beta_0 + s(cigs, 4.6) + s(faminc, 2.9) + parity + male + white + \epsilon \quad 7)$$



Obrázek 1. Vztahy mezi závisle proměnnou a jednotlivými regresory po odhadu rovnice 6.

4.2 Out-of-sample

Pro porovnání predikční schopnosti byly využity odhady získané z předešlé části, které byly aplikovány na out-of-sample data. Pro porovnání obou předpovědí byl použit Diebold-Mariano test, jehož hodnota vyšla -0,095 což odpovídá p-value 0.32. Na základě této hodnoty nelze zamítnout nulovou hypotézu o shodnosti obou předpovědí.

5 ZÁVĚR

V práci byla porovnána predikční schopnost klasického lineárního modelu s modelem GAM, kde jako bazické funkce byly použity smoothing splines. Po naučení obou modelů na in-the-sample vzorku a vybrání statisticky významných proměnných, byly oba odhady aplikovány na out-of-sample vzorek. Rozdíl v predikci byl pak vyhodnocen na základě Diebold – Mariano testu. Z analýzy vyšlo, že GAM model, který obsahoval dvě proměnné, jejichž vývoj byl aproximován pomocí smoothing spline funkce, neposkytuje statisticky lepší predikci, než klasický lineární model. Přestože na obrázku 1 byl vidět nelineární charakter mezi váhou

narozeného dítěte a množstvím vykouřených cigaret, respektive příjmem rodiny, nebyla tato nelinearita natolik významná, aby klasický lineární model nebyl schopný poskytnout statisticky stejné výsledky jako GAM model.

Použitá literatura

1. CASELLA, G., FIENBERG, S., OLKIN, I. 2013. An Introduction to Statistical Learning with application in R. Springer New York Heidelberg Dordrecht London
2. DIEBOLD, F. X., MARIANO, R.S. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13(3), 253–263.
3. HASTIEM, T., TIBSHIRAMI, R. 1986. Generalized additive models. *Statistical Science* 1, 297–318.
4. HASTIEM, T., TIBSHIRAMI, R. 1990. *Generalized Additive Models*. (New York: Chapman and Hall).
5. HE, X., LAURS, M. 1998. Bycatch, finning, and economic value of blue shark in the Hawaii-based longline fishery. In: *Shark Management and Conservation. Proceedings from the Sharks and Man Workshop of the Second World Fisheries Congress (Synopsis)*. Department of Primary Industries, Queensland, pp. 88–94.
6. HE, X., BIGELOW, K.A., BOGGS, C.H. 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. *Fish. Res.*, 31 (1997), pp. 147–158
7. WOOLDRIDGE, J. M. 2006. *Introductory econometrics: a modern approach*. Mason, OH, Thomson/South-Western

Kontaktné údaje

Ing. Lukáš Frýd

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky

nám. Winstona Churchilla 1938/4,

Praha 3 130 67

Email: lukas.fryd@gmail.com

SOFTVÉROVÉ NÁSTROJE NA PRIESTOROVÚ ANALÝZU DÁT

SOFTWARE TOOLS FOR SPATIAL DATA ANALYSIS

Andrea Furková, Michaela Chocholatá

Abstrakt

Empirické práce zamerané na priestorovú štatistiku a ekonometriu značne závisia od dostupnosti špecializovaného softvéru. Dlhé obdobie priestorová analýza dát trpela nedostatkom vhodných softvérových nástrojov. V súčasnej dobe si vedci môžu vybrať z pomerne širokej škály programov, tzv. toolbox nástrojov alebo tzv. open source programov. Príspevok prináša stručný prehľad softvéru, ktorý je v súčasnosti k dispozícii na priestorovú analýzu údajov. Väčšia pozornosť v článku bude venovaná softvéru GeoDa, ktorý je jedným z najpopulárnejších softvérových možností pre priestorovú analýzu údajov.

***Kľúčové slová:** priestorová analýza dát, GeoDa, R program*

Abstract

Empirical evidence related to the spatial statistical and econometric topics highly depends on availability of specialized software. For many years spatial data analysis suffered from the lack of proper software tools. Nowadays scientists can chose from relatively wide range of programs, toolboxes or open source efforts. This paper provides a brief software overview currently available for implementing a spatial data analysis. More attention will be paid to software GeoDa which is one of the most popular software for spatial data analysis.

***Keywords:** spatial data analysis, GeoDa, R program*

1 INTRODUCTION

The implementation of spatial statistical and econometric methods in empirical work in regional science and urban economics for many years has suffered from the lack of specialized software. The lack of such tools was criticized in the late 1980s by Haining (1989) and cited as a major barrier to the adoption and use of spatial statistics and spatial econometrics by researchers. Researchers from many environmental sciences such as geography, geology, climatology, oceanography and also from scientific fields such economy or statistics highlighted geographic information as an inevitable part of their scientific research. Why do they need to know this information? Because everything that happens, happens in our street, in our town or city, region, state, country and continent. Geography is vitally important in our everyday lives and understanding the geography around us and its importance is essential to our everyday lives. Also the first Law of Geography, according to Tobler, states that "everything is related to everything else, but near things are more related than distant things" (see Tobler, 1970, p. 234). Initially attention related to the development of specialized software for spatial data analysis tented to focus on conceptual issues, such as how to integrate spatial statistical methods and a GIS (Geographic Information Systems or Science) environment and which techniques would be most advisable in such a framework. In recent years, the situation has changed dramatically and a number of options for applying spatial data analysis are currently on offer to the interested research. Possibilities start from

programs, scripts and extensions for statistical and GIS packages, to open source effort using software environments such as R, Java or Python. The purpose of this paper is twofold. First, it focuses on a brief software overview currently available for implementing a spatial data analysis and next in more details we will deal with GeoDa (Geographic Data Analysis) package which is one of the most popular software for spatial data analysis. Our aim is to describe main functions, utilities and techniques supported by GeoDa (<https://geodacenter.asu.edu/software/downloads>).

The paper is organized as follows. The second part provides a brief overview of software possibilities for spatial data analysis, the third part presents GeoDa software package and finally main conclusions are the part of the last section of this paper.

2 SOFTWARE OVERVIEW FOR SPATIAL DATA ANALYSIS

Historically, the first opportunity available for researchers engaged in spatial data analysis is represented by the SpaceStat package (Anselin, 1992a; 1992b) available as freestanding program. The first version dates back to 1991 and it has been upgraded several times. Main features offered by SpaceStat regard the possibility of inputting and manipulating spatial weights matrices (for more details see e.g. Smith, 2014), some features relating to Explanatory Spatial Data Analysis – ESDA (for more details see e.g. Bivand, 2010b). The program allows the Maximum Likelihood (ML) estimation of the parameters in the *spatial lag* and the *spatial error* models as well as the testing of the various misspecification hypotheses, including normality (the Jarque-Bera test), heteroscedasticity (the Breusch-Pagan test) and spatial dependence (the Lagrange Multiplier and Likelihood Ratio test). The main drawback of the SpaceStat is an unpleasant and old-fashioned interface based on the “Gauss” language.

A dramatic evolution with respect to SpaceStat represents GeoDa which was developed by Luc Anselin and his co-workers (Anselin, 2003, 2004; Anselin et al., 2006). As a main advantage over SpaceStat is considered to be a user-friendly graphical interface based on a windows environment. GeoDa will be presented in more details in next chapter therefore now we go on with a brief description of the R for spatial data analysis. In general, R is a free, open-source, object oriented language and environment for statistical computing and graphics. Free and open-source means that anyone is free to use, redistribute and change the software in any way. Using R is worthwhile especially it is free and user does not only get to use the functions that are built in the software but user can create his own. In R, the fundamental unit of shareable code is the package which bundles together code, data, documentation, and tests. These packages can be easily shared with others. On April 2016, there were over 8 200 packages available on the Comprehensive R Archive Network - CRAN (see http://www.econ.uiuc.edu/~lab/workshop/Spatial_in_R.html). Installation of R versions and R packages are available at <http://cran.r-project.org>. R includes many functions that can be used for reading, visualising, and analysing spatial data. The packages concerning spatial topics can be roughly structured into the few main topics listed in the Table1. In general spatial data comes in many “shapes” and “sizes” and the most common types of spatial data are points, lines, polygons or grid. Spatial data usually comes in so called shapefile (.shp file) which stores non topological geometry and attribute information for the spatial features in a data set. This information is a starting point for spatial data analysis in any specialized software for spatial data analysis (for more details see e.g. Bivand, 2010a). Manipulating with spatial data in R is conditioned by installation of the proper packages. Other R packages cover many

functions to create spatial weight matrix object from polygon contiguities, from point patterns by distance and tessellations, for summarising these objects and for permitting their use in spatial data analysis. Many tools of ESDA are accessible and many different hypotheses can be tested through the R packages. As we have already pointed out, advanced user can incorporate his own functions or create a whole package. It is also necessary to mention that packages listed in Table 1 are only a small fraction of all available packages specialized in spatial topics (see <http://cran.r-project.org> and <https://CRAN.R-project.org/view=Spatial>).

Topic	Packages
Classes for spatial data	<i>sp, spdep, rgdal, rgeos, raster, shp2graph</i>
Handling spatial data	<i>raster, gdistance, GeoXp, geosphere, trip</i>
Reading and writing spatial data	<i>rgdal, gdal, maps, maptools, shapefiles, OpenStreetMaps</i>
Visualisation	<i>RColorBrewer, mapview, tmap, quickmapr, leaflet</i>
Point pattern analysis	<i>spatial, spatsta, spatgraphs, ecspace</i>
Geostatistics	<i>gstat, geoR, GeoRglm, gstat</i>
Spatial regression	<i>nlme, spdep, shet, McSpatial, splm</i>

Table 1: Main R packages for spatial data analysis

Source: own elaboration

Besides the R language and environment, nowadays a number of initiatives set up by individual researchers or groups have made specific routines for spatial data analysis available. Perhaps, the best known are toolboxes for spatial econometric analysis in software such as Matlab or Stata (for more details see Arbia, 2006).

3 A SPATIAL DATA ANALYSIS IN GEODA

GeoDa is a relatively recent free program designed to implement techniques for exploratory data analysis on spatial data in the form of points or polygons in a geographical space. It represents a dramatic evolution with respect to SpaceStat because of user-friendly graphical interface. Also in comparison with R environment from this point of view, GeoDa surpluses because GeoDa is completely driven by a point and click interface and does not require any programming. On the other hand, in terms of the range of spatial statistical techniques included, GeoDa is most alike to the collection of functions developed in the R environment. The contents of *sp* and *spdep* packages are very similar to tools incorporated in GeoDa. GeoDa is not customizable or extensible by the user, which is one of the strengths of the R. In that sense, GeoDa and R are seen as highly complementary, ideally with more sophisticated users ‘graduating’ to R after being introduced to the techniques in GeoDa.

Category	Function
Spatial data	Data input from shape file (point, polygon) Data input from text (to point or polygon shape) Data output to text (data or shape file) Create grid polygon shape file from text input Centroid computation Thiessen polygons
Data transformation	Variable transformation (log, exp, etc.) Queries, dummy variables (regime variables) Variable algebra (addition, multiplication, etc.) Spatial lag variable construction Rate calculation and rate smoothing Data table join
Mapping	Generic quantile choropleth map Standard deviational map Percentile map Outlier map (box map) Circular cartogram Map movie Conditional maps Smoothed rate map Excess rate map
Exploratory Data Analysis	Histogram Box plot Scatter plot Parallel coordinate plot Three-dimensional scatter plot Conditional plot (histogram, box plot, scatter plot)
Spatial autocorrelation	Spatial weights creation (Rook, Queen, distance, k -nearest) Higher order spatial weights Spatial weights characteristics (connectedness histogram) Moran scatterplot with inference Bivariate Moran scatterplot with inference Moran scatterplot for rates (EB standardization) Local Moran significance map Local Moran cluster map Bivariate local Moran Local Moran for rates
Spatial regression	OLS with diagnostics (e.g., LM test, Moran's I) Maximum likelihood spatial lag model Maximum likelihood spatial error model Predicted value map Residual map

Table 2: GeoDa functionality overview

Source: own elaboration based on Anselin et al. (2010)

GeoDa design is interactive that combines maps with statistical graphs, using the technology of dynamically linked windows. GeoDa offers only the analysis of discrete geospatial data, i.e. objects characterized by their location in space either as points (point coordinates) or

polygons (polygon boundary coordinates). A shape file discussed before is a standard for storing spatial information. GeoDa contains functionality to read and write such files, as well as to convert ASCII text input files for point coordinates or boundary file coordinates to the shape file format. The software implementation of GeoDa consists of two components: the user interface and graphics windows (based on Microsoft Foundation Classes - MFC) on the one hand and the computational engine (including statistical operations, randomization, spatial regression is pure C++ code and largely cross platform based on) on the other hand. GeoDa is mainly oriented at graphical tools, simple descriptive spatial analysis such as spatial autocorrelation statistics, and the analysis of spatial outliers and a wide range of functions relating to ESDA. The spatial econometric regression routines in GeoDa are limited in estimation of the classical linear regression via OLS and ML estimation of the parameters associated with the *spatial error* and the *spatial lag* models (for more details see e.g. Anselin et al., 2010). The basic diagnostics for spatial dependence, spatial heteroskedasticity and normality are available for the standard OLS regression residuals. Asymptotic inference is based on the Likelihood Ratio Test and on an estimate of the asymptotic covariance matrix using the algorithm developed by Smirnov (for more details see Arbia, 2006). GeoDa enables the construction of different types of spatial weight matrices reading a shape file. However, the estimation procedures only permit symmetric structures for the spatial weights (e. g. contiguity or distance-based weights) and they cannot be performed on more sophisticated structures such as e.g. *k*-nearest neighbours weighting scheme.

The functionality served by GeoDa can be classified into six categories. The categories and associated functions are given in Table 2. The latest version GeoDa 1.8 contains several new features such as support for more spatial file formats, full space-time data support in all views, basemap layers for all maps, averages charts, scatter plot matrices, nonparametric spatial autocorrelation (correlogram), and flexible data categorization.

4 CONCLUSION

An important factor that helped promote the dissemination of spatial data analysis and spatial econometric methods to empirical practice was the availability of a growing number of software tools. Nowadays, researcher can choose among relatively large number of software tools. This paper was aimed at main software tools overview for spatial data analysis, more detailed in GeoDa software. Our intention was to provide brief useful survey to reader interested in spatial topics.

References

1. ANSELIN, L. 1992a. SpaceStat: A Program for the Analysis of Spatial Data. National Centre for Geographic Information and Analysis, University of California.
2. ANSELIN, L. 1992b. SpaceStat Tutorial: A Workbook for Using SpaceStat in the Analysis of Spatial Data. National Centre for Geographic Information and Analysis, University of California.
3. ANSELIN, L. 2003. GeoDa 0.9 User's Guide. Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois.
4. ANSELIN, L., 2004. GeoDa 0.95 Release Notes. Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois.
5. ANSELIN, L. SYABRI, I. and KHO, Y. 2006. GeoDa: an Introduction to spatial data analysis. In: Geographical Analysis, 38(1), pp. 5-22.

6. ANSELIN, L., SYABRI, I. and KHO, Y. 2010. GeoDa: an introduction to spatial data analysis. In: Handbook of applied spatial analysis. Software Tools, Methods and Applications. Springer Berlin Heidelberg, pp. 73-89.
7. ARBIA, G. 2006. Spatial econometrics: statistical foundations and applications to regional convergence. Springer Science & Business Media, pp. 207.
8. BIVAND, S. 2010a. Spatial econometric functions in R. In: Handbook of applied spatial analysis. Software Tools, Methods and Applications. Springer Berlin Heidelberg, pp. 53-71.
9. BIVAND, S. 2010b. Exploratory Spatial Data Analysis. In: Handbook of applied spatial analysis. Software Tools, Methods and Applications. Springer Berlin Heidelberg, pp. 219-254.
10. HAINING, R. 1989. Geography and spatial statistics: current positions, future developments. In: Remodelling Geography. Basil Blackwell, Oxford, pp.191-203.
11. SMITH, E. T. 2014. Spatial Weight Matrices. Available at: http://www.seas.upenn.edu/~ese502/lab-content/extra_materials/SPATIAL%20WEIGHT%20MATRICES.pdf [Accessed 09/26/2014].
12. TOBLER, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography, 46(2), pp. 234 – 240.
13. <http://cran.r-project.org> [Accessed 09/26/2016].
14. <https://CRAN.R-project.org/view=Spatial> [Accessed 09/26/2016].
15. http://www.econ.uiuc.edu/~lab/workshop/Spatial_in_R.html [Accessed 09/26/2016].
16. <https://geodacenter.asu.edu/software/downloads> [Accessed 05/02/2015].

Acknowledgements

This work was supported by the Grant Agency of Slovak Republic – VEGA grant No. 1/0285/14 “Regional modelling of the economic growth of EU countries with concentration on spatial econometric methods”.

Authors' addresses

Ing. Andrea Furková, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 832

Email: andrea.furkova@euba.sk

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

VLIV ZAOKROUHLOVACÍ CHYBY NA REALIZOVANÝ ROTPTYL

IMPACT OF ROUNDING ERROR ON REALIZED VARIANCE

Vladimír Holý

Abstrakt

Při odhadu integrované variance cen finančních aktiv je nutné zohlednit mikrostrukturní šum, který značně vychyluje realizovaný rozptyl. Jedna ze složek mikrostrukturního šumu je chyba způsobená zaokrouhlováním cen. Analyzujeme měnový pár EUR/CZK a zkoumáme rozptyl zaokrouhlovacího šumu a jeho vliv na celkové vychýlení realizovaného rozptylu.

Klíčová slova: *Vysokofrekvenční data, Zaokrouhlovací chyba, Mikrostrukturní šum, Realizovaný rozptyl, Forex*

Abstract

When estimating the integrated variation of financial asset prices, it is necessary to take into account the microstructure noise which makes realized variance biased. One of the components of the microstructure noise is an error caused by rounding of prices. We analyze the currency pair EUR/CZK and investigate the variance of rounding noise and its impact on the overall bias of realized variance.

Keywords: *High-Frequency Data, Rounding Error, Microstructure Noise, Realized Variance, Forex*

1 ÚVOD

Dostupnost finančních vysokofrekvenčních dat nabízí nové možnosti a výzvy v odhadu rozptylu cen akcií a měnových kurzů. Používání vyšších frekvencí přináší více informací, ale v datech se pak objevuje tzv. mikrostrukturní šum, který značně vychyluje realizovaný rozptyl. Tento šum je způsoben mnoha faktory: bid-ask spreadem, asymetrickou informací, chybami v zaznamenávání cen a v neposlední řadě i zaokrouhlováním cen. V tomto článku se budeme zabývat komponentou mikrostrukturního šumu, která představuje chybu vzniklou zaokrouhlováním cen.

2 TEORETICKÉ ZÁKLADY

Nechť p_τ^* je jednorozměrný proces logaritmických cen a předpokládejme, že se jedná o spojitý semimartingal

$$p_\tau^* = p_0^* + \int_0^\tau \mu(s) ds + \int_0^\tau \sigma(s) dW_s, \quad (1)$$

kde $\mu(s)$ je càglàg drift proces, $\sigma(s)$ je càglàg volatility proces a W_s značí Wienerův proces. Proces p_τ^* se nazývá eficientní cena. Místo p_τ^* ovšem ve skutečnosti pozorujeme diskrétní proces p_{t_i} definovaný jako

$$p_{t_i} = p_{t_i}^* + \varepsilon_{t_i}, \quad (2)$$

kde ε_{t_i} je mikrostrukturní šum. V literatuře je toto nejběžnější situace, jak je uvedeno např. v knize [3] nebo v knize [1]. Pro odhad rizika aktiva se používá tzv. integrovaná variance (IV)

definovaná jako

$$IV(a, b) = \int_a^b \sigma^2(s) ds. \quad (3)$$

Přirozeným estimátorem IV je tzv. realizovaný rozptyl (RV) daný předpisem

$$RV^n = \sum_{i=1}^n (p_{t_i} - p_{t_{i-1}})^2. \quad (4)$$

Pokud uvažujeme proces kontaminovaný mikrostrukturním šumem (2), můžeme rozepsat RV jako

$$RV^n = \sum_{i=1}^n (p_{t_i}^* - p_{t_{i-1}}^*)^2 + 2 \sum_{i=1}^n (p_{t_i}^* - p_{t_{i-1}}^*) (\varepsilon_{t_i} - \varepsilon_{t_{i-1}}) + \sum_{i=1}^n (\varepsilon_{t_i} - \varepsilon_{t_{i-1}})^2. \quad (5)$$

Druhý a třetí člen v (5) způsobí, že při velkém počtu pozorování je RV značně vychýlený.

3 ŠUM ZPŮSOBENÝ ZAOKROUHLOVÁNÍM

Mikrostrukturní šum přítomný v (2) rozložíme na

$$\varepsilon_{t_i} = \varepsilon_{t_i}^O + \varepsilon_{t_i}^R, \quad (6)$$

kde $\varepsilon_{t_i}^R$ je šum způsobený zaokrouhlováním a $\varepsilon_{t_i}^O$ je všechen ostatní šum způsobený dalšími faktory. Zaokrouhlovací šum $\varepsilon_{t_i}^R$ je deterministická funkce náhodných veličin $p_{t_i}^*$ a $\varepsilon_{t_i}^O$, kterou lze zapsat jako

$$\varepsilon_{t_i}^R = f(p_{t_i}^*, \varepsilon_{t_i}^O) = \lfloor p_{t_i}^* + \varepsilon_{t_i}^O \rfloor - p_{t_i}^* - \varepsilon_{t_i}^O, \quad (7)$$

kde symbolem $\lfloor x \rfloor$ značíme zaokrouhlení čísla x na nejbližší celé číslo. Pokud je pozorovatelná cena p_{t_i} zaokrouhlena na β desetinných míst, zaokrouhlovací šum nabývá hodnot z intervalu

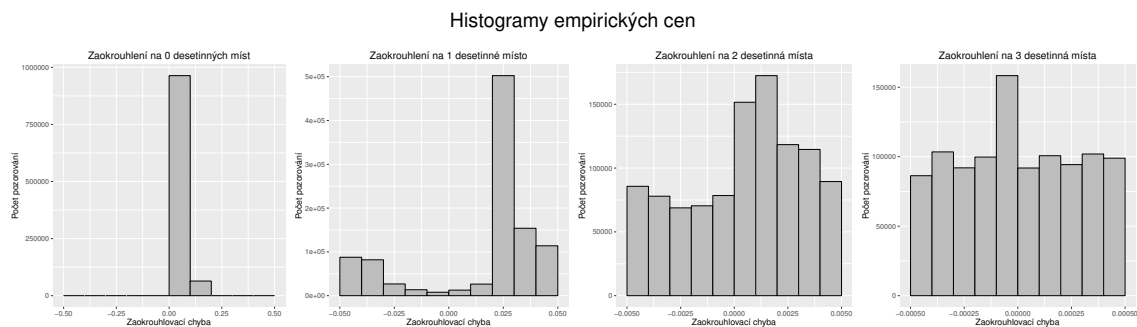
$$\varepsilon_{t_i}^R \in (0.5 \cdot 10^{-\beta}, 0.5 \cdot 10^{-\beta}). \quad (8)$$

Pokud pro jednoduchost budeme uvažovat nulovou sériovou korelaci šumu $\varepsilon_{t_i}^R$ a nulovou korelaci s eficientní cenou $p_{t_i}^*$ a šumem $\varepsilon_{t_i}^O$, lze vychýlení realizovaného rozptylu způsobené zaokrouhlením vypočítat jako

$$\sum_{i=1}^n (\varepsilon_{t_i}^R - \varepsilon_{t_{i-1}}^R)^2 = 2n \cdot \text{var}(\varepsilon_{t_i}^R). \quad (9)$$

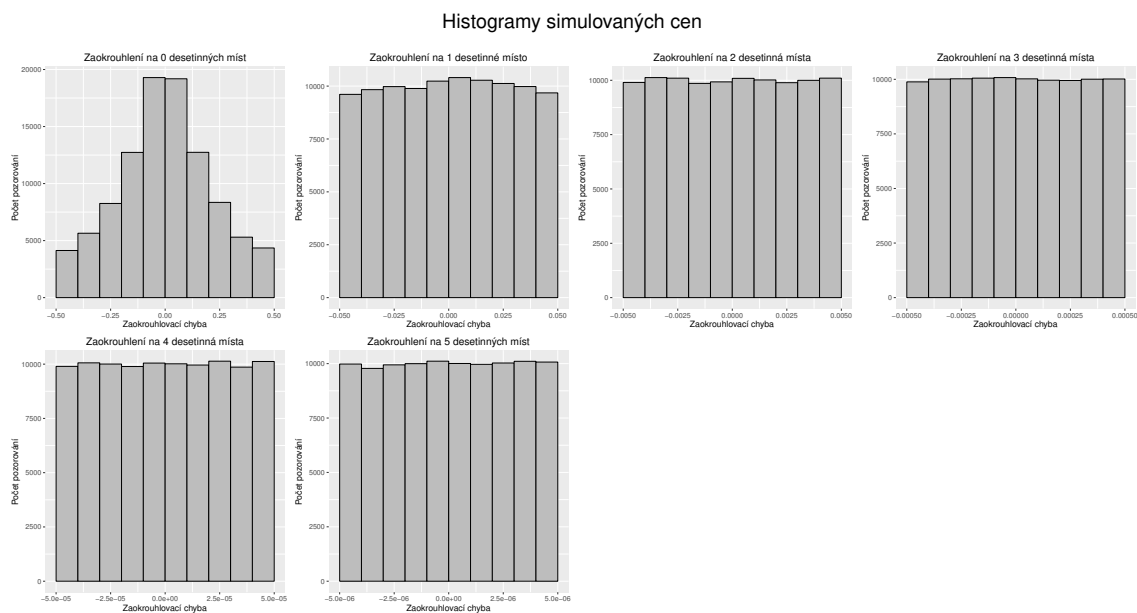
4 ANALÝZA ROZDĚLENÍ ZAOKROUHLOVACÍCH CHYB

V této sekci budeme zkoumat distribuční funkci zaokrouhlovacího šumu $\varepsilon_{t_i}^R$. Analyzujeme časovou řadu měnového páru EUR/CZK od 1.1.2016 do 31.10.2016. Ceny toho měnového páru jsou v datech zaokrouhleny na 4 desetinná místa. Pokud data zaokrouhlíme na méně desetinných míst, můžeme zkoumat empirické rozdělení odchylek od původních cen zaokrouhlených na 4 desetinná místa. V obrázku 1 jsou vykresleny histogramy empirického rozdělení cen při zaokrouhlování na 0 až 3 desetinná místa.



Obrázek 1: Histogramy zaokrouhlovacích chyb empirických cen.

Podobným způsobem můžeme v grafu 2 zobrazit histogramy pro ceny získané ze simulací. Proces p_{t_i} uvažujeme jako geometrický Brownův pohyb a provedeme 10^5 simulací. Empirické a zejména simulované histogramy naznačují, že při zaokrouhlení na vyšší počet desetinných míst se rozdělení šumu p_{t_i} podobá rovnoměrnému rozdělení.



Obrázek 2: Histogramy zaokrouhlovacích chyb simulovaných cen.

Podobnost se spojitým rovnoměrným rozdělením na intervalu $(-0.5 \cdot 10^{-\beta}, 0.5 \cdot 10^{-\beta})$ se pokusíme ověřit statistickými testy. Použijeme Pearsonův χ^2 test na četnosti prvních nenulových číslic v šumu $\varepsilon_{t_i}^R$. Pro simulované ceny ještě použijeme Kolmogorovův-Smirnovův test pro celý šum $\varepsilon_{t_i}^R$. Výsledky v tabulce 1 ukazují, že pro empirický šum hypotézu o rovnoměrně rozdělených číslicích zamítáme. To ovšem může být způsobeno zaokrouhlením na malý počet číslic. V případě simulovaného šumu oba testy potvrdili rovnoměrné rozdělení při zaokrouhlení na 2 číslice a více.

Desetinná místa	0	1	2	3	4	5
Empirická p-hodnota χ^2	0.000	0.000	0.000	0.000	-	-
Simulovaná p-hodnota χ^2	0.000	0.000	0.076	0.749	0.687	0.667
Simulovaná p-hodnota K-S	0.000	0.000	0.697	0.793	0.661	0.199

Tabulka 1: P-hodnoty Pearsonova χ^2 testu pro první číslici a Kolmogorova-Smirnovova testu.

5 ANALÝZA ROZPTYLU ZAOKROUHLOVACÍCH CHYB

Více než samotné rozdělení šumu $\varepsilon_{t_i}^R$ nás zajímá pouze jeho rozptyl. Pokud předpokládáme, že šum $\varepsilon_{t_i}^R$ zaokrouhlený na β desetinných míst je rovnoměrně rozdělený, lze jeho rozptyl spočítat jako

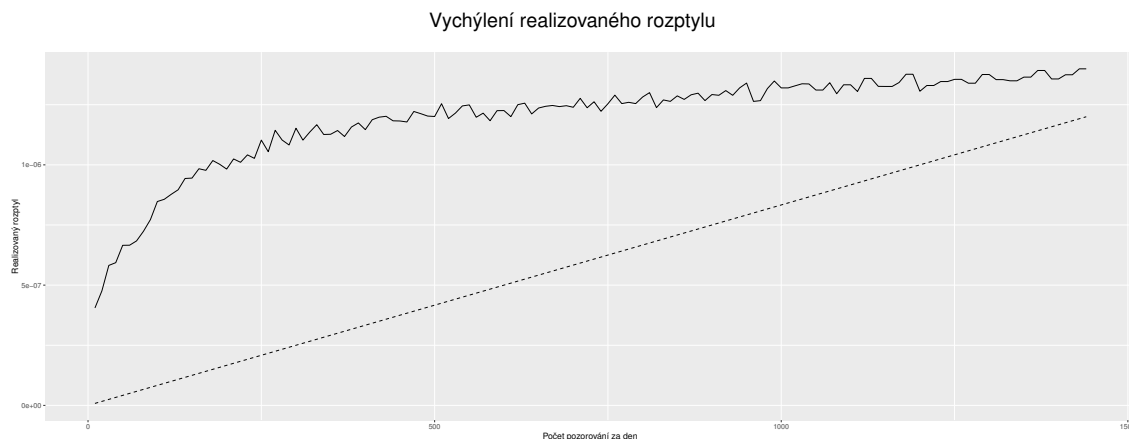
$$\text{var}(\varepsilon_{t_i}^R) = \frac{1}{12} 10^{-\beta}. \quad (10)$$

V tabulce 2 jsou uvedeny výběrové rozptyly pro empirické i simulované ceny a rozptyly při předpokladu rovnoměrného rozdělení. Při zaokrouhlení na vyšší počet číslic se výběrové rozptyly velmi podobají rozptylu rovnoměrného rozdělení.

Desetinná místa	0	1	2	3	4	5
Empirický rozptyl	$9.30 \cdot 10^{-4}$	$7.66 \cdot 10^{-4}$	$7.28 \cdot 10^{-6}$	$8.62 \cdot 10^{-8}$	-	-
Simulovaný rozptyl	$4.75 \cdot 10^{-2}$	$8.17 \cdot 10^{-4}$	$8.35 \cdot 10^{-6}$	$8.35 \cdot 10^{-8}$	$8.33 \cdot 10^{-10}$	$8.33 \cdot 10^{-12}$
Rozptyl rovn. roz.	$8.33 \cdot 10^{-2}$	$8.33 \cdot 10^{-4}$	$8.33 \cdot 10^{-6}$	$8.33 \cdot 10^{-8}$	$8.33 \cdot 10^{-10}$	$8.33 \cdot 10^{-12}$

Tabulka 2: Rozptyly empirických a simulovaných cen a rozptyly rovnoměrného rozdělení.

Tyto výsledky nás tedy motivují pro modelování rozptylu šumu $\varepsilon_{t_i}^R$ podle vzorce (10). Graf 3 nám ukazuje průměrné denní vychýlení realizovaného rozptylu při různých počtech pozorování pro měnový pár EUR/CZK. Spojitá čára představuje pozorované vychýlení způsobené celkovým šumem ε_{t_i} , přerušovaná čára pak představuje vychýlení způsobené zaokrouhlováním modelované podle vzorců (9) a (10). Vychýlení způsobené zaokrouhlováním je zřejmě lineární, zatímco celkové vychýlení má složitější strukturu. Ostatní složky šumu ε_{t_i} totiž mohou být závislé v čase a také závislé na eficientní ceně. V reálných datech se často objevuje negativní korelace s eficientní cenou, což má za následek zápornost celkového vychýlení pro větší počet pozorování, jak je popsáno v článku [2].



Obrázek 3: Celkové vychýlení realizovaného rozptylu (nepřerušovaná čára) a vychýlení způsobené pouze zaokrouhlováním (přerušovaná čára).

6 ZÁVĚR

Empirická a simulační studie ukázala, že při zaokrouhlování cen na vyšší počet číslic je možné počítat rozptyl zaokrouhlovacího šumu pomocí vzorce pro rozptyl rovnoměrného rozdělení. Analýza cen měnového páru EUR/CZK, které jsou k dispozici zaokrouhlené na 4 desetinná místa, pak odhalila, že zaokrouhlovací šum je významnou součástí celkového mikrostrukturního šumu.

Poděkování

Tento článek byl napsán v rámci grantu IGS F4/63/2016 Vysoké školy ekonomické v Praze.

Použitá literatura

- [1] AIT-SAHALIA, Y. a JACOD, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press. ISBN 9788578110796. doi: 10.1017/CBO9781107415324.004.
- [2] HANSEN, P. R. a LUNDE, A. (2006). Realized Variance and Market Microstructure Noise. *Journal of Business & Economic Statistics*, **24**(2), 127–161. ISSN 0735-0015. URL <http://www.jstor.org/stable/27638860>.
- [3] HAUTSCH, N. (2011). *Econometrics of financial high frequency data*. Springer Science & Business Media. ISBN 9783642219245.

Kontaktní údaje

Ing. Tomáš Novák, PhD.
 Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky
 Dolnozemska 1/b, 852 35 Bratislava
 Tel: (421) 267 295 824
 email: novak@euba.sk

NÁVRH ROBUSTNÍHO ZÁCHRANNÉHO SYSTÉMU POMOCÍ MAXIMÁLNÍHO SCÉNÁŘE

ROBUSTNESS OF EMERGENCY SYSTEM DESIGN BY MAX-HULL SCENARIO

Jaroslav Janáček, Marek Kvet

Abstrakt

Návrh záchranného obslužného systému spočívá v umístění omezeného počtu obslužných středisek v některých z možných umístění tak, aby byly uspokojeny požadavky uživatelů na obsluhu a aby bylo minimalizováno zatížení průměrného uživatele. O zatížení uživatele je předpokládáno, že je úměrné jeho vzdálenosti od nejbližšího obslužného střediska. Návrh robustního systému je definován jako návrh, který bere v úvahu dané scénáře tak, že minimalizuje maximální hodnotu účelové funkce optimální pro jednotlivé scénáře. Min-max vazební podmínky používané v standardních modelech představují nežádoucí zátěž každého solveru úloh celočíselného programování kvůli špatné konvergenci metody větví a hranic. V tomto článku se snažíme tento nedostatek překonat konstrukcí a použitím takzvaného maximálního scénáře.

***Klíčové slova:** návrh záchranného systému, radiální formulace, robustnost, maximální scénář*

Abstract

Emergency service system design locates a limited number of service centers at some of possible locations to satisfy the users' demands for service with the objective to minimize disutility perceived by an average user. The disutility is assumed to be proportional to the distance of a user location from the nearest service center. A robust service system design is defined as the design, which complies with specified scenarios by minimizing the maximal objective function of the individual instances corresponding with the particular scenarios. The min-max link-up constraints used in standard models represent an undesirable burden of any IP-solver due to bad convergence of the branch-and-bound method. Within this paper, we try to overcome the drawback following from the link-up constraints by constructing and using so-called max-hull scenario.

***Keywords:** emergency system design, radial formulation, robustness, max-hull scenario*

1 INTRODUCTION

When an emergency service system is designed, the designer must take into account that traversing time between a service center and the affected user might be impacted by various random events following weather or traffic. Thus, the system designer must face the demand for system resistance to such critical events [1], [6]. Most of the approaches to increasing the system resistance are based on making its design resistant to possible detrimental scenarios, which can appear in the road network as a consequence of random events. An individual scenario is characterized by particular time distances between a users' location and a possible service center location. A robust service system design has to comply with all the given scenarios. The usual way of taking into account all scenarios is based on minimizing the maximal objective function of the individual instances corresponding with the particular scenarios. The min-max link-up constraints represent an undesirable burden in any integer

programming solving tool. An emergency service system design for a given road network in serviced area locates limited number of service centers at positions from a given set of possible locations to satisfy the system users' demands. Within this paper, the amount of perceived disutility is assumed to be proportional to the distance from a user location to the nearest located service center and then, sum of time distances from particular system users to the nearest located service center is minimized [2], [4], [5]. According to [1], we focus on the min-sum emergency service system design, which is robust considering given finite set of scenarios. To model the design problem, a radial formulation with homogenous system of radii [3] is employed. We present the original approach [1] to the robust design using the radial formulation with premature termination of the branch-and-bound method [6] and compare it to the approach using the so-called max-hull scenario.

The remainder of the paper is organized as follows: Section 2 is devoted to the description of the original robust design of the emergency service system with min-sum objective including the radial formulation. The max-hull scenario construction and usage are described in Section 3 and the associated numerical experiments are performed and compared in Section 4. The findings are summarized in Section 5.

2 ROBUST EMERGENCY SYSTEM DESIGN USING SCENARIO SET

The robust emergency system design problem with homogeneous radial formulation can be described according to [3], [4] and [6] in the following way. Let symbols J and I denote the set of users' locations and the set of possible service center locations respectively. Symbol b_j denotes the number of users, who share the location j and p denotes the maximal number of service centers, which can be chosen from I . The objective is to minimize the maximal system disutility over the set U of scenarios by suitable deployment of p centers. The objective function value under a given scenario is defined as the sum of all disutility values associated with the individual users. Disutility following from the distance between locations i and j under a scenario $u \in U$ is denoted as d_{iju} . Within this paper, we consider that each value of d_{iju} is integer and less than or equal to the value $v+1$. The strategic decisions are modeled by decision variables $y_i \in \{0, 1\}$, $i \in I$. The variable y_i takes the value of 1 if a service center is located at i and it takes the value of 0 otherwise. The variable h is used as the upper bound of the objective function values connected with individual scenarios.

In addition, zero-one variables x_{jsu} for $s = 0 \dots v$ and $u \in U$ are introduced to express disutility perceived by an individual user. The variable x_{jsu} takes the value of 1, if the disutility under the scenario $u \in U$ of a user at $j \in J$ from the nearest located center is greater than s and it takes the value of 0 otherwise. Then the expression $x_{j0u} + x_{j1u} + x_{j2u} + \dots + x_{jvu}$ expresses the value of disutility d_{ju^*} from the user location j to the nearest located service center under the scenario $u \in U$. We introduce a zero-one constant a_{iju}^s under the scenario $u \in U$ for $i \in I, j \in J$ and $s \in [0..v]$. The constant a_{iju}^s is equal to 1, if the disutility d_{iju} between the user location j and the possible center location i is less or equal to s , otherwise a_{iju}^s is equal to 0. Then, the radial-type min-sum robust emergency service system design problem can be formulated as follows [6].

$$\text{Minimize } h \tag{1}$$

$$\text{Subject to: } x_{jsu} + \sum_{i \in I} a_{iju}^s y_i \geq 1 \quad \text{for } j \in J, \quad s = 0, 1, \dots, v, \quad u \in U \tag{2}$$

$$\sum_{i \in I} y_i \leq p \tag{3}$$

$$\sum_{j \in J} b_j \sum_{s=0}^v x_{jsu} \leq h \text{ for } u \in U \quad (4)$$

$$y_i \in \{0, 1\} \text{ for } i \in I \quad (5)$$

$$x_{jsu} \geq 0 \text{ for } j \in J, s = 0, 1, \dots, v, u \in U \quad (6)$$

$$h \geq 0 \quad (7)$$

In this model, the objective function (1) represented by variable h gives the upper bound of the all objective function values over the set U of scenarios. The constraints (2) ensure that the variables x_{jsu} are allowed to take the value of 0, if at least one center is located in radius s from the user location j and the constraint (3) causes that the number of located service centers does not exceed p . The link-up constraints (4) ensure that each individual objective function value is less than or equal to the upper bound h . The constraints (5), (6) and (7) are usual obligatory constraints.

3 ROBUST EMERGENCY SYSTEM DESIGN USING MAX-HULL SCENARIO

By inspection of the model (1)-(7), we can find that the subscripts i and u form inseparable pair in all variables and all constraints excepting variables y_i and link-up constraints (4). It follows that each size of the model (either number of variables or number of constraints) grows proportionally with the cardinality of the set of scenarios. The growing size of the model may cause technical insolvability of the problem with big number of scenarios. That is why we concentrate on suggesting such approximate methods, which can replace the above approach by another one, which does not depend on the number of scenarios. One of the possible suggestions is to inspect all the scenarios for each pair (i, j) , where $i \in I$ and $j \in J$ and determine d_{ij-Max} according to (8).

$$d_{ij-Max} = \max\{d_{iju} \mid u \in U\} \quad (8)$$

Then, after introducing auxiliary variables x_{js} for $s = 0 \dots v$ and $j \in J$ and the zero-one constants a_{ij}^s for $i \in I, j \in J$ and $s \in [0..v]$ analogically to variables x_{jsu} and constants a_{iju}^s , the following model can be constituted instead the former model (1)-(7).

$$\text{Minimize } \sum_{j \in J} b_j \sum_{s=0}^v x_{js} \quad (9)$$

$$\text{Subject to: } x_{js} + \sum_{i \in I} a_{ij}^s y_i \geq 1 \text{ for } j \in J, s = 0, 1, \dots, v \quad (10)$$

$$\sum_{i \in I} y_i \leq p \quad (11)$$

$$y_i \in \{0, 1\} \text{ for } i \in I \quad (12)$$

$$x_{js} \geq 0 \text{ for } j \in J, s = 0, 1, \dots, v \quad (13)$$

4 NUMERICAL EXPERIMENTS

Within this section, we present the overview of numerical experiments aimed at comparison of the results obtained by earlier and later suggested approaches to robust service system

design. To solve the problems described in the previous sections, the optimization software FICO Xpress 7.9 (64-bit, release 2015) was used and the experiments were run on a PC equipped with the Intel® Core™ i7 5500U processor with the parameters: 2.4 GHz and 16 GB RAM. The used benchmarks were derived from the real emergency health care system, which was originally implemented in selected regions of Slovakia. For each self-governing region, i.e. Bratislava (BA), Banská Bystrica (BB), Košice (KE), Nitra (NR), Trenčín (TN), Trnava (TT) and Žilina (ZA), all cities and villages with corresponding number of inhabitants b_j were taken in hundreds. The set of communities represents both the set J of users' locations and also the set I of possible service center locations. The cardinalities of these sets vary from 87 to 515 according to the considered region. The number p of located centers was derived from the original design and it varies from 9 to 52. To enlarge the set of benchmarks, we added also several instances for the region of Žilina with different values of parameter p . The network distance from a user to the nearest located service center was taken as the user's disutility. The achieved results are summarized in the Table 1. Due to the lack of common benchmarks for study of robustness, ten detrimental scenarios used in our computational study were generated randomly in the way described in [6].

An individual experiment was organized so that the referential system design was suggested for the basic scenario only without considering the other scenarios. The particular results for the basic scenario are reported in the "BASIC" part of the table. The optimal solution for the basic scenario b is denoted by \mathbf{y}^b and the objective function corresponding with the basic scenario is computed according to (14), where b is substituted for a general scenario u .

$$f^u(\mathbf{y}) = \sum_{j \in J} b_j \min\{d_{iju} : i \in I, y_i = 1\} \quad (14)$$

Let us denote \mathbf{y}^r the robust solution obtained by solving the problem (1)-(7), where the optimization process was limited to 15 minutes in this study. Furthermore, \mathbf{y}^M denotes the solution obtained by solving the problem (9)-(13). The corresponding results are reported in the parts "STANDARD ROBUST" and "MAX-HULL" of the table respectively. To evaluate any obtained solution from the point of the worst case, objective function $f^w(\mathbf{y})$ is introduced according to (15).

$$f^w(\mathbf{y}) = \max\{f^u(\mathbf{y}) : u \in U\} \quad (15)$$

For sake of brevity, we report the values of objective functions in percentage, where the value of $f^b(\mathbf{y}^b)$ is considered to be basis. The corresponding denotation of the relative values for a solution \mathbf{y} is $F^u(\mathbf{y})$, $F^b(\mathbf{y})$ or $F^w(\mathbf{y})$ in Table 1. Symbol "CT" denotes the computational time in seconds.

	I	p	BASIC		STANDARD ROBUST			MAX-HULL		
			CT	$F^w(\mathbf{y}^b)$	CT	$F^b(\mathbf{y}^r)$	$F^w(\mathbf{y}^r)$	CT	$F^b(\mathbf{y}^M)$	$F^w(\mathbf{y}^M)$
BA	87	9	0.1	170.23	17.9	108.40	124.95	0.1	120.33	133.63
BB	515	52	1.2	110.22	902.4	102.44	107.57	5.4	108.62	111.49
KE	460	46	2.1	112.89	912.2	102.29	106.33	1.3	112.19	113.71
NR	350	35	1.2	114.36	902.9	102.52	106.83	1.6	109.51	110.98
TN	276	28	3.6	123.03	902.4	104.05	111.77	5.5	113.74	116.17
TT	249	25	0.6	115.11	803.4	103.87	108.93	1.9	114.60	117.42
ZA	315	32	0.7	124.15	900.6	103.75	109.57	0.7	109.48	113.40
Žilina	315	158	0.9	108.39	169.2	100.29	103.72	1.1	105.52	106.75
	315	105	1.3	108.01	209.7	101.90	105.69	2.3	109.14	111.98
	315	79	1.8	115.94	237.4	101.59	107.76	1.8	116.25	118.46
	315	63	1.4	117.52	901.8	106.63	111.49	0.7	118.06	122.27
	315	32	0.7	124.15	900.6	103.75	109.57	0.7	109.48	113.40
	315	21	0.8	121.13	902.0	109.68	116.73	3.4	115.68	120.90
	315	16	1.2	128.94	902.5	103.50	120.61	1.4	115.51	122.94

Table 1: The results of numerical experiments for the self-governing regions of Slovakia

5 CONCLUSIONS

This paper was focused on mastering real-sized instances of the emergency service system design problem under uncertainty using commercial IP-solver. The uncertainty follows from the requirement of making the designed system resistant to various detrimental scenarios. Within this paper, we compared the robust approach based on scenarios with the suggested max-hull approach. The achieved results have confirmed that the max-hull approach enables to get the results faster and it is acceptable also from the viewpoint of solution accuracy. Based on performed computational study, we can conclude that we have presented a useful tool for the robust emergency service system design, which can be easily implemented using common commercial optimization software.

Acknowledgement

This work was supported by the research grants VEGA 1/0518/15 “Resilient rescue systems with uncertain accessibility of service”, APVV-15-0179 “Reliability of emergency systems on infrastructure with uncertain functionality of critical elements” and by the project University Science Park of the University of Žilina (ITMS: 26220220184) supported by the Research & Development Operational Program funded by the European Regional Development Fund.

References

1. CORREIA, I., SALDANHA da GAMA, F. 2015. Facility locations under uncertainty. Location Science, eds. Laporte, Nickel, Saldanha da Gama, pp. 177-203
2. CURRENT, J., DASKIN, M., SCHILLING, D. 2002. Discrete network location models. In: Drezner Z. et al.: Facility location. Applications and theory. Berlin: Springer, pp. 81-118
3. JANÁČEK, J., KVET, M. 2012. Relevant Network Distances for Approximate Approach to Large p-Median Problems. In: Operations research proceedings 2012, September 4-7, 2012, Hannover, Germany, pp. 123-128
4. JANÁČEK, J., KVET, M. 2013. Public service system design with disutility relevance estimation. In: Proceedings of the 31st international conference “Mathematical Methods in Economics”, September 11-13, 2013, Jihlava, Czech Republic, pp. 332-337
5. JANÁČEK, J., KVET, M. 2015. Emergency system design with temporarily failing centers. In: SOR 2015: proceedings of the 13th international symposium on operational research, Ljubljana: Slovenian Society Informatika, Section for Operational Research, pp. 490-495
6. JANÁČEK, J., KVET, M. 2016. Designing a Robust Emergency Service System by Lagrangean Relaxation. In: Proceedings of the 34th international conference “Mathematical Methods in Economics”, September 6-9, 2016, Liberec: Technical University of Liberec, pp. 349-353

Kontaktné údaje

Prof. RNDr. Jaroslav Janáček, CSc.- Ing.Marek Kvet, PhD

Žilinská univerzita v Žiline, Fakulta riadenia a informatiky

Univerzitná 8215/1, 010 26 Žilina

Tel: (421) 041/513 4204

email: jaroslav.janacek@fri.uniza.sk – marek.kvet@fri.uniza.sk

POJEDNANIE O EFEKTÍVNUM POČTE STANÍC ZÁCHRANNEJ ZDRAVOTNEJ SLUŽBY

DISQUISITION ON THE EFFECTIVE NUMBER OF THE EMERGENCY MEDICAL SERVICE STATIONS

Marta Janáčková, Alžbeta Szendreyová

Abstrakt

Cieľom nášho príspevku je overiť hypotézu, ktorá navrhuje vhodný počet staníc záchranej zdravotnej služby (ZZS) v jednotlivých krajoch Slovenskej republiky. Pre zlepšenie dostupnosti zásahu ZZS je výhodnejší väčší počet staníc ZZS. Ich počet nie je možné zvyšovať ľubovoľne, najmä z ekonomických dôvodov. Výsledky, ktoré dostaneme na základe hypotézy porovnáme s počtom staníc reálne umiestnených v jednotlivých krajoch SR v súčasnosti.

Príučové slová: verejný obslužný systém, p -medián, záchranná zdravotná služba

Abstract

The aim of our paper is to test the hypothesis that suggests the suitable number of the Emergency medical services stations (EMS) in different regions of Slovakia. To improve the availability of EMS intervention is preferable to have a larger number of EMS stations. Their number cannot be increased arbitrarily, mainly for economic reasons. The results, that we get on the hypothesis, we compare with the real number of stations located in different regions of Slovakia today.

Keywords: public service systems, p -median problem, emergency medical services

1 ÚVOD

Optimalizačné úlohy o umiestnení staníc ZZS sa obyčajne zaoberajú vhodným výberom (lokalizáciou) umiestnenia vopred určeného počtu staníc zo všetkých možných umiestnení. Teraz sa budeme zaoberať vhodným počtom umiestňovaných staníc. Vhodný počet je daný dostupnosťou služby pre zákazníka. Dostupnosť služby je ovplyvnená predovšetkým cestnou sieťou, demografickým zložením obyvateľstva, predpismi o poskytovaní záchranej zdravotnej služby a pod. Pri nových výberových konaniach sa počet staníc mení len minimálne. Zistíme, do akej miery výsledky nami overovanej hypotézy zodpovedajú súčasnému stavu, prípadne či súčasný existujúci počet staníc ZZS je v zmysle hypotézy vyhovujúci.

2 HYPOTÉZA

Optimálne umiestnenie p staníc je možné riešiť ako úlohu p -mediánu. Pre náš príspevok formulujeme úlohu nasledovne:

Majme danú cestnú sieť a maticu najkratších vzdialeností medzi jednotlivými uzlami siete. Z uzlov siete vyberieme množinu I prípustných kandidátov na umiestnenie strediska a množinu J všetkých zákazníkov. Cieľom úlohy je umiestniť daný počet p stredísk do

niektorých uzlov $i \in I$ a z nich obslúžiť každý uzol $j \in J$ tak, aby súčet vzdialeností medzi priradenými uzlami i, j bol minimálny.

Je zrejmé, že zväčšením čísla p v úlohe o p -mediáne dostaneme optimálne riešenie s menšou hodnotou účelovej funkcie (UF). Ak budeme úlohu o p -mediáne riešiť na intervale $\langle p_{min}, p_{max} \rangle$ bez protiváhy, ktorá by tlačila proti umiestneniu stanice, potom optimálny výsledok dostaneme pre p_{max} . Predbežné experimenty ukázali, že závislosť hodnoty účelovej funkcie optimálneho riešenia úlohy o p -mediáne na počte umiestnených staníc p bude mať rýdzo konvexný a klesajúci charakter. Priemerný pokles hodnoty UF pre dve po sebe idúce p je na intervale $\langle p_{min}, p_{max} \rangle$ určený vzťahom $(UF(p_{min}) - UF(p_{max})) / (p_{max} - p_{min})$. Vyslovíme nasledujúcu hypotézu:

Efektívny počet umiestnených staníc nastane v takom p_{ef} , kde pokles hodnoty UF pri zvýšení p o jednotku bude rovný alebo blízky hodnote priemerného poklesu UF na danom intervale.

3 EXPERIMENTY

Hodnoty p_{min} a p_{max} majú výrazný vplyv na výpočet priemerného poklesu hodnôt účelovej funkcie. Pre tento príspevok sme sa rozhodli použiť krajné možnosti, t. j. $p_{min} = 1$ a $p_{max} = |I|$. Výpočty robíme na cestnej sieti SR. Pre každý kraj SR riešime sadu úloh o p -mediáne pre rastúce p z intervalu $\langle p_{min}, p_{max} \rangle$. Z hodnôt UF z optimálnych riešení a príslušných intervalov pre daný kraj vypočítame v každej sade priemerný pokles hodnoty UF a tiež aktuálny pokles pre každé dve po sebe idúce hodnoty p . Výber z výsledkov optimálnych riešení p -mediánu pre rastúce p v Trnavskom kraji je uvedený v tabuľke 1.

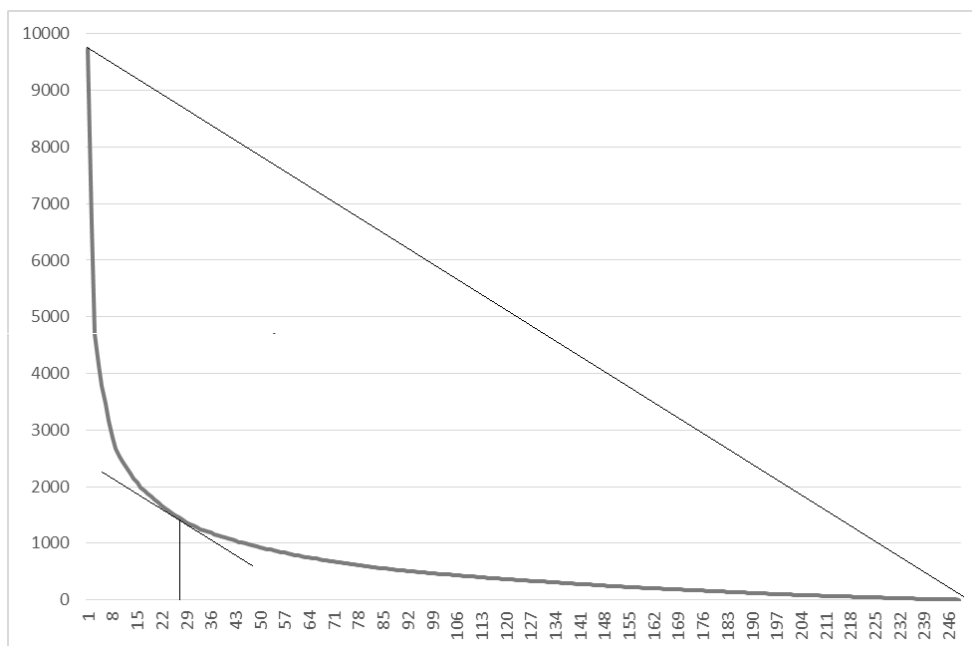
Tabuľka 1: Trnavský kraj, 249 obcí, $p_{min} = 1$, $p_{max} = 249$, smernica = 39,254

P	HUF	Pokles	P	HUF	Pokles	P	HUF	Pokles
1	9735	0	15	2064	75	29	1374	38
2	6641	3094	16	1989	75	30	1336	38
3	4692	1949	17	1925	64	31	1305	31
4	4141	551	18	1870	55	32	1276	29
5	3786	355	19	1817	53	33	1249	27
6	3466	320	20	1765	52	34	1226	23
7	3169	297	21	1714	51	35	1203	23
8	2860	309	22	1664	50	36	1181	22
9	2678	182	23	1617	47	37	1159	22
10	2536	142	24	1573	44	38	1138	21
11	2437	99	25	1530	43	39	1118	20
12	2331	106	26	1490	40	40	1098	20
13	2233	98	27	1450	40	41	1079	19
14	2139	94	28	1412	38	42	1060	19

Priemerný pokles hodnoty UF pre daný interval predstavuje smernicu priamky, ktorá umožní nájsť p_{ef} .

V každej sade nájdeme také (jedno alebo viac) p_{ef} , v ktorom aktuálny pokles hodnoty UF najlepšie aproximuje hodnotu priemerného poklesu UF pre danú sadu. Názorný príklad

závislosti hodnôt UF na počte umiestnení stredísk p z výsledkov optimálnych riešení pre Trnavský kraj uvádzame v grafe 1.



Graf 1: Závislosť UF na hodnote p pre Trnavský kraj

Z priebehu závislosti je zrejmé, že pre p menšie ako p_{ef} je pokles aktuálnej hodnoty UF väčší ako priemerný pokles pre daný interval. Opačne, pre p väčšie ako p_{ef} je pokles medzi susednými hodnotami UF menší ako priemerný pokles pre daný interval, t.j. pridanie ďalšieho strediska je v zmysle zlepšenia hodnoty UF menej efektívne. Preto takto určené p_{ef} nazývame efektívnou hodnotou pre odhad vhodného počtu na umiestnenie stredísk. Porovnanie hodnôt p_{ef} , získaných na základe uvedenej hypotézy s počtom uzlov, v ktorých sú v súčasnosti umiestnené stanice ZZS, uvádzame v tabuľke 2.

Tabuľka 2: Porovnanie vypočítaného a reálneho p

Kraj	p_{max}	HUF ($p = 1$)	Priemerný pokles	Interval p_{ef}	Reálny počet staníc
BA	87	2227	25,895	13-15	15
BB	515	28453	55,356	40-43	36
KE	460	24263	52,861	33-36	32
NR	350	13236	37,926	34-36	27
PO	664	35691	53,833	46-50	32
TN	276	10388	37,775	28-30	22
TT	249	9735	39,254	25-30	17
ZA	315	14356	45,72	26-30	29

4 ZÁVER

Na základe údajov uvedených v tabuľke 2 môžeme povedať, že počet reálne umiestnených staníc v jednotlivých krajoch SR je prevažne menší ako vypočítané hodnoty p_{ef} . V zmysle

uvedenej hypotézy a použitých parametrov sú teda existujúce počty staníc zvolené efektívne. Je zrejmé, že vzhľadom na účel pôsobnosti staníc ZZS je umiestnenie 1-2 staníc, ale aj naopak napríklad 80 a viac staníc v kraji nereálne. Núka sa otázka, ako vhodne určiť krajné hodnoty p_{min} a p_{max} , a aký vplyv budú mať tieto zmeny na určenie efektívneho počtu stredísk p_{ef} . Týmito otázkami sa chceme zaoberať neskôr.

Acknowledgements

This work was supported by the research grants VEGA 1/0518/15 “Resilient rescue systems with uncertain accessibility of service” and APVV-15-0179 “Reliability of emergency systems on infrastructure with uncertain functionality of critical elements”.

Použitá literatúra

1. JANÁČEK, J. a kol, Navrhovanie územne rozľahlých obslužných systémov. 1. vyd., Žilina : Žilinská univerzita, 2010, 404 s.
2. JANÁČKOVÁ, M., SZENDREYOVÁ, A. An importance of the population density for the location of the Emergency Medical Service stations. In: Mathematical methods in economics: 34th international conference, Liberec, Czech Republic, September 6th-9th, 2016, s. 354-358.
3. JANAČKOVÁ, M., SZENDREYOVÁ, A. Time-distance versus utility in the public service system design. In: SOR '15 Proceedings of the 13th International Symposium on Operational Research. Bled, Slovenia, September 23-25, 2015. Ljubljana, Section for Operational Research, s. 446-451.

Kontaktné údaje

RNDr. Marta Janáčková, CSc.
Žilinská univerzita v Žiline, Fakulta strojnícka
Univerzitná 1, 010 26 Žilina
Tel: (421 41) 513 4953
email: marta.janackova@fstroj.uniza.sk.

RNDr. Alžbeta Szendreyová, PhD.
Žilinská univerzita v Žiline, Fakulta riadenia a informatiky
Univerzitná 1, 010 26 Žilina
Tel: (421 41) 513 4205
email: alzbeta.szendreyova@fri.uniza.sk

TEORETICKÉ VÝCHODISKÁ MODELU S KOREKČNÝM ČLENOM¹

THEORETICAL BACKGROUND OF AN ERROR CORRECTION MODEL

Miroslava Jánošová

Abstrakt

Cieľom predkladaného príspevku je priblíženie teoretických východísk modelu s korekčným členom. Ide o modely patriace do skupiny modelov vektorovej autoregresie, ktoré využívajú testovanie stacionarity časových radov. Práve tieto modely sú v súčasnosti často využívaným modelovým aparátom pre ich schopnosť popísania tak krátkodobých ako aj dlhodobých vzťahov. Dlhodobá dynamika je zabezpečená prostredníctvom zakomponovania pôvodných, nestacionárnych časových radov a krátkodobá dynamika je zachytená v diferencovaných, stacionárnych časových radoch.

KLúčové slová: *stacionarita časových radov, model s korekčným členom, diferencia časových radov*

Abstract

The aim of the following paper is to provide the theoretical background of an error correction model. These models belong to the group of vector autoregressive models which rely on stationarity tests. Particularly these models are nowadays a frequently applied modelling framework due to their capability to describe short-term as well as long-term relationships. Long-term dynamics is kept by incorporating original, non-stationary time series and short-term dynamics is captured in differentiated, stationary time series.

Keywords: *stationary time series, error correction model, differentiated time series*

ÚVOD

Modely s korekčným členom je možné zaradiť do skupiny modelov vektorovej autoregresie s testovaním stacionarity. Na časové rady, ktoré vykazujú nestacionaritu je využívaná analýza kointegrácie, ktorá postupne viedla k vzniku modelov s korekčným členom. Využívaním redukcie stupňa oneskorenia mal za následok vznik jednoduchších tvarov modelov. Modely s korekčným členom sú charakteristické ich snahou zachytiť tak krátkodobé dynamické vzťahy ako aj dlhodobé rovnovážne vzťahy.

Tieto modely, ktoré predstavujú modely zo skupiny vektorovej autoregresie skúmajú prepojenie časových radov prostredníctvom závislosti medzi posunutými hodnotami. Práve týmto je v modeloch zabezpečená previazanosť časových radov. Tieto modely sú špecifickou formou viacrovnicových modelov, kde jednotlivé vysvetľované premenné závisia od svojich oneskorených hodnôt a od oneskorených hodnôt všetkých ostatných premenných, ktoré v modeli vystupujú.

Nevyhnutným predpokladom konštrukcie takéhoto modelu je preskúmanie stacionarity využívaných časových radov. Informácie o stacionarite získavame prostredníctvom diferencií

¹ Tento 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)“.

a podľa počtu diferencií. Práve so stacionaritou sa často krát stretávame pri časových radoch makroekonomických a finančných dátach, a práve vtedy je potrebné použiť diferenci príslušného radu. Konceptia, ktorá je zameraná na danú problematiku bola navrhnutá Engle a Grangerom. V nasledujúcom texte si bližšie popíšeme koncepciu modelu s korekčným členom.

1 MODEL S KOREKČNÝM ČLENOM

Dynamiku krátkodobých nerovnovážnych vzťahov medzi premennými môžeme vyjadriť prostredníctvom už spomínaného modelu s korekčným členom. Momentálne ide o jeden z najčastejšie využívaných modelov.

Tento typ modelu je využívaný na určenie miery vybočenia systému od tzv. ekvilibria. Pri konštrukcii modelu daného typu modelu sa vychádza z konštrukcie modelu VAR². Predpokladáme l - rozmerný model VAR(p) typu I(1) v tvare:

$$X_t = c + \Omega D_t + \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + a_t \quad (1)$$

pričom $\{a_t\}$ predstavuje l - rozmerný gaussovský proces bieleho šumu s kovariančnou maticou Σ_a , c sú konštanty, D_t je deterministickou zložkou. Ak model VAR zapíšeme v nasledujúcej forme:

$$\Delta X_t = c + \Omega D_t + \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_p \Delta X_{t-p-1} + \Pi X_{t-p} a_t \quad (2)$$

Pričom

$$\Gamma_i = -(I_l - \Phi_1 - \dots - \Phi_i) \text{ pre } i=1, \dots, p-1, \Pi = -(I_l - \Phi_1 - \dots - \Phi_p) \quad (3)$$

uvedený model je modelom korekcie chyby. Jednotlivé vzťahy a to dlhodobé ako aj krátkodobé sú vyjadrené prostredníctvom parametrickej matice Π .

V modeli môže prísť k trom nasledujúcim vzťahom³:

1) $h(\Pi) = l$, tj. matica Π má plnú hodnotu, čo znamená, že l - rozmerná časová rada je generovaná stacionárnym vektorovým procesom $\{X_t\}$ takýto prípad môže nastať pre $p = 2, \Phi_2 = 0_{l \times l}$ a $\Phi_1 = (\rho_1, \rho_2, \dots, \rho_l) I_l$ kde $|\rho_i| < 1$ pre $i = 1, \dots, l$. Táto situácia je logická, v prípade, že viacrozmerná časová rada je stacionárna, v prípade, že sú stacionárne jednotlivé časové rady k ich stacionarizácii nie je potrebné diferencovanie. Ak sa aj napriek tomu prikróčí k diferencovaniu je potrebné aby model diferencii bol vytvorený na princípe modelu VAR a bol prítomný aspoň jeden nediferencovaný člen.

2) $h(\Pi) = 0$, to znamená, že matica Π je nulová a model (2) neobsahuje nediferencovaný člen, l -rozmerná časová rada je generovaná nestacionárnym vektorovým procesom $\{X_t\}$, jej stacionarizáciu je možné vykonať prostredníctvom individuálneho diferencovania jednotlivých časových rad, takýto prípad nastáva napr. keď $p = 2, \Phi_2 = 0_{l \times l}$ a $\Phi_1 = I_l$. Pri diferencovaní nedochádza k strate informácií o dlhodobom vzťahu medzi časovými radami vzhľadom na skutočnosť, že žiaden takýto vzťah neexistuje.

² Vektorovo autoregresný proces kde, model VAR(p) môžeme zapísať v tvare

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + a_t$$

³ ARLT, J – ARLTOVÁ, M. 2007. Ekonomické časové rady. Praha: Garda Publishing, 2007. s. 238

3) $0 < h(\Pi) = r < 1$, takomto prípade nie je odstránený nediferencovaný člen modelu (2), ale súčasne nie je možné považovať vektorový proces $\{X_t\}$ za stacionárny. Vzhľadom na to, že matica Π je nenulová je možné medzi časovými radami nájsť dlhodobý vzťah a stacionarizáciu vykonať bez straty informácie prostredníctvom individuálneho diferencovania. To znamená, že niektoré časové rady je možné stacionarizovať prostredníctvom diferencovania preto, lebo nie sú súčasťou žiadneho dlhodobého vzťahu medzi inými časovými radami. Iné časové rady však nie je možné stacionarizovať preto, lebo ich lineárna kombinácia s inými časovými radami je už stacionárna a tieto rady sú kointegrované.

V literatúre však častejšie nachádzame modely s korekčným členom, ktoré sú zapísané s autoregresne rozdelenými oneskoreniami – model ADL⁴.

Model s korekčným členom, prostredníctvom, ktorého je možné vyjadrenie tak dlhodobého ako aj krátkodobého vzťahu medzi dvomi premennými je možné zapísať v tvare:

$$\Delta y_t = a_0 + b_0 \Delta x_t + (a_1 - 1)(y_{t-1} - \delta_1 x_{t-1}) + u_t \quad (4)$$

pričom $\delta_1 = -(b_0 + b_1)/(a_1 - 1)$.

Krátkodobú rovnováhu vo vzťahu predstavuje $\Delta y_t = a_0 + b_0 \Delta x_t + u_t$ a korekčný člen resp. chybu ekvilibria vyjadruje vzťah $(a_1 - 1)(y_{t-1} - \delta_1 x_{t-1})$.

Parameter a_1 interpretujeme ako mieru zotrvačnosti, pričom, čím je táto zotrvačnosť väčšia, tým je menší vplyv nerovnováhy a teda $(1 - a_1)$ sa približuje k nule zľava. Aby model (4) mal dlhodobé vzťahy, musí byť zachovaná platnosť $|a_1| < 1$, v prípade porušenia platnosti podmienky dochádza k narušeniu mechanizmu korekčného člena a v takom prípade parameter znázorňujúci korekčného člena je kladný. Ak je hodnota koeficienta a_1 blízka 1 ide o rýchlu konvergenciu k rovnováhe pri približovaní sa k 0 ide o pomalú konvergenciu k rovnováhe.

Vo vzťahu vystupuje parameter b_0 , ktorý predstavuje krátkodobý multiplikátor a indikuje nám vplyv krátkodobej závislosti medzi vysvetľovanou a vysvetľujúcou premennou (krátkodobú elasticitu).

Parameter δ_1 je dlhodobým multiplikátorom a vyjadrujeme prostredníctvom neho dlhodobú elasticitu vysvetľovanej premennej vo vzťahu k vysvetľujúcej premennej.

Vzťah $\gamma = (a_1 - 1)$ predstavuje koeficient krátkodobého prispôsobenia. Tento koeficient by mal dosahovať záporné hodnoty a meria rýchlosť prispôsobenia k rovnováhe. Pri hodnotách blížiacich sa 1 ide o rýchle prispôsobovanie sa k rovnováhe naopak pri hodnotách blížiacich sa 0 ide o pomalé prispôsobovanie sa rovnováhe.

Pri modeloch na princípe člena korigujúceho chybu nevzniká problém falošnej regresie, vzhľadom na skutočnosť, že premenné Δy_t a Δx_t sú stacionárnym a y_t a x_t sú kointegrované. Ich lineárna kombinácia $(y_{t-1} - \delta_1 x_{t-1})$ je stacionárny. Pri odhade parametrov modelu ECM je možné využitie metódy najmenších štvorcov a pri hodnotení odhadov t-test a F-test štatistickej významnosti.

Model založený na princípe EC je možné rovnako použiť aj pre viac nezávislých premenných, a teda budeme vychádzať z modelu ADL(1,1,p)

$$Y_t = a_0 + a_1 Y_{t-1} + b_{01} X_{t1} + \dots + b_{0p} X_{tp} + b_{11} X_{t-1,1} + \dots + b_{1p} X_{t-1,p} + u_t \quad (5)$$

po matematickej úprave dostávame všeobecný model s korekčným členom:

⁴ Autoregressive Distributed Lags (1,1) – model, ktorý predpokladá časové oneskorenie o jedno obdobie, zapíšeme v tvare $y_t = a_0 + a_1 y_{t-1} + b_0 x_t + b_1 x_{t-1} + u_t$ kde u_t spĺňa podmienky bieleho šumu.

$$\Delta Y_t = a_0 + b_{01}\Delta X_{t1} + \dots + b_{0p}\Delta X_{tp} - (1 - a_1)(Y_{t-1} - X_{t-1,1} - \dots - X_{t-1,p}) + (a_1 - 1 + b_{01} + b_{11})X_{t-1,1} + \dots + (a_1 + b_{0p} + b_{1p})X_{t-1,p} + u_t \quad (6)$$

Vzťah (6) je mechanizmom prispôsobenia sa závislej premennej k dlhodobej rovnováhe prostredníctvom pôsobenia všetkých vysvetľujúcich premenných.

ZÁVER

V príspevku sa venujeme teoretickému prístupu modelu korekcie chyby, ktorého konštrukcia vychádza z VAR modelu. Charakterizovali sme tri základné situácie, ktoré môžu vzniknúť v dôsledku tvorby modelu s korekčným členom. Následne sme poukázali aj na modely s korekčným členom, ktoré sú zapísané s autoregresne rozdelenými oneskoreniami. Hlavnou motiváciou vzniku predkladaného príspevku bolo preskúmanie možností ako pristupovať k modelovaniu ak je našim cieľom zobrazenie tak krátkodobých ako aj dlhodobých vzťahov medzi modelovanými premennými. Práve modely založené na princípe korekcie chyby túto možnosť poskytujú.

Použitá literatúra

1. ARTL, J. - ARTLOVÁ, M.: Ekonomické časové rady. Praha: Grada Publishing, 2007. ISBN 978-80-247-1319-9.
2. ENGLE, R. F. – GRANGER, C. W. J.: Co-integration and error correction representation, estimation and testing. In.: *Econometrica* 55, 1988, s. 251- 276.
3. LUKÁČIK, M. – PEKÁR, J.: Kointegračná analýza v ekonometrii. Dostupné na internete: <http://www.fhi.sk/files/katedry/kove/veda-vyskum/prace/2006/Lukacik-Pekar2006.pdf>

Kontaktné údaje

Ing. Miroslava Jánošová
 Ekonomický ústav Slovenskej akadémie vied
 Šancová č. 56
 811 05 Bratislava
 email: miroslava.dolinajcova@savba.sk

EFEKTÍVNOSŤ REGIONÁLNIHO PREROZDELENIA ZDROJOV EÚ NA ÚROVNI NUTS 3 NA SLOVENSKU V ROKOCH 2007-15

EFFECTIVITY OF REGIONAL ALOCATION OF EU SOURCES AT NUTS 3 LEVEL IN SLOVAKIA DURING 2007-15 PERIOD

Ivan Lichner, Marek Radvanský, Tomáš Miklošovič

Institute of Economic Research SAS¹

Abstrakt

Slovensko v rámci programového obdobia 2007-2013 implementovalo zo Štrukturálnych a Kohézneho fondu prostriedky v objeme 11,5 mld. Eur. Cieľ implementácie bol definovaný v NSRR sa sústredil len na dosiahnutie konvergencie k priemeru EÚ na úrovni krajiny a nie z hľadiska možnosti ovplyvnenia regionálnej konvergencie v rámci krajiny. Predložený článok sa venuje porovnaniu objemu implementácie v rámci jednotlivých krajov SR na úrovni NUTS 3, ako aj identifikácií možnej efektívnosti takto vynaložených prostriedkov.

Kľúčové slová: Kohézny fond, štrukturálne fondy, model HERMIN, regionálna politika

Abstract

During the programming period 2007-2013 Slovakia implemented 11.5 bil. EUR from Structural and Cohesion Funds. Aims of EU funds implementation were defined in NSRF and focused only on convergence of Slovak regions to level of EU average. Convergence of Slovak regions among each other was not taken into account. This paper represents assessment of regional distribution of EU funds and identification of possible effectiveness of those funds.

Keywords: Cohesion fund, Structural funds, HERMIN model, regional policy

ÚVOD

Predpoklady vplyvu politiky súdržnosti v prípravnej fáze programového obdobia 2007-2013 boli zhrnuté v rámcovom dokumente „Ex-ante hodnotenie Národného strategického referenčného rámca“ vytvoreného Ekonomickým ústavom SAV v roku 2006 (Šikula a kol., 2006). V rámci uvedeného hodnotenia sa konštatuje, že strategická časť dokumentu venujúca sa očakávaným vplyvom politiky súdržnosti je orientovaná národohospodársky so zreteľom na dosiahnutie celkovej konvergencie SR ku krajinám EÚ bez dostatočného zreteľa na regionálny rozmer, ktorý mal byť následne dosiahnutý skôr individuálne v plánoch jednotlivých operačných programov. V zhrnutí hodnotenia sa na strane 20 píše: „Navrhované kontextové ukazovatele, ktoré majú hodnotiť realizáciu podpory sú často formulované tak, že nebude možné ich vykazovať na úrovni NUTS 2 a nižšie, teda regionálny rozmer týchto intervencií nebude možné preukázať.“ Keďže ciele implementácie ani objem alokácií

¹ Tento článok bol podporený Agentúrou na Podporu Výskumu a Vývoja na základe zmluvy k projektu č. SK-PL-2015-0058.

v detailnejšej miere nezohľadňovali potrebu vnútornej kohézie ani potrebu relatívne rovnomernejšieho rozdelenia prostriedkov medzi krajinami SR, hlavným zámerom predloženého článku je porovnanie reálneho a relatívneho objemu implementovaných prostriedkov vzhľadom na ekonomickú výkonnosť kraja, ako aj odhadnutie dosiahnutej (očakávanej) efektívnosti takto vynaložených prostriedkov.

1. METODIKA

Na analýzu vplyvov politiky súdržnosti sa používajú viaceré kvalitatívne a kvantitatívne metódy, ktorých súčasťou sú aj modelové simulácie a hodnotenia založené na ekonometrických modeloch. Základnou motiváciou analýz dopadov ŠF a KF je získanie odpovedí na otázky, ktoré sú priamo prepojené na efektívnosť politiky súdržnosti v redukovaní regionálnych disparít, efektívnosti distribúcie finančných prostriedkov a potrebu reformy politiky súdržnosti v prípade prístupu ďalších krajín do EÚ. Prvými snahami o hodnotenie politiky súdržnosti pomocou ekonometrického modelu boli modelové simulácie využívajúce model HERMES. Tento model však bol použitý iba pre Írsko. Nasledovníkom modelu HERMES bol model HERMIN, ktorý už pokryl väčší geografický priestor v rámci EÚ. Začiatkom deväťdesiatych rokov bol model HERMIN použitý na hodnotenie vplyvov ŠF a KF v Írsku a Portugalsku, neskôr pribudli Grécko a Španielsko. V priebehu prvej dekády nového milénia bol model HERMIN Európskou Komisiou (EK) vytvorený pre všetky krajiny EÚ 27. Okrem modelu HERMIN využila EK na analýzu politiky súdržnosti aj model QUEST (Varga a in't Veld, 2009), ktorý predstavuje dopytovo orientovaný typ modelu. V súčasnosti sa zvyšuje snaha o nájdenie nasledovníka týchto modelov a vznikajú rôzne alternatívne typy modelových prístupov (MASST, EuImpactMod, RHOMOLO, GMR-EUROPE a pod.), ktoré sa však zatiaľ zápasia s problematikou ich aplikovateľnosti v praxi. Väčšina funkčných regionálnych modelov v Európe používa regionálne členenie na úrovni NUTS 2. Vzhľadom na decentralizáciu, ktorá na Slovensku prebehla v roku 1996 však ako funkčné regióny môžeme označiť úroveň NUTS 3, t. j. samosprávne kraje. Z tohto dôvodu bol autormi upravený regionálny model HERMIN na úroveň NUTS 3 regiónov. Detailná štruktúra modelu je opísaná v Radvanský a kol. (2016).

Počiatky modelu HERMIN možno hľadať v komplexnom, viac sektorovom modeli HERMES, ktorý bol vyvinutý Európskou komisiou na začiatku roka 1980 (d' Alcantara a Italianer, 1982). HERMIN bol pôvodne navrhnutý, ako menšia verzia modelu HERMES, aby bola aplikovateľná v podmienkach, kde je problém s dostupnosťou údajov, napr. v chudobnejších, menej rozvinutých členských štátoch a regiónoch EÚ na západnej a južnej periférii (Írsko, Severné Írsko, Portugalsko, Španielsko, talianskej Mezzogiorno, a Grécko). Dôsledkom nedostatku podrobných údajov a dostatočne dlhých časových radov, bez štruktúrnych zmien bolo, že modelový rámec musel byť založený na pomerne jednoduchom teoretickom rámci. Táto relatívna jednoduchosť je práve jednou z hlavných výhod modelu HERMIN.

Jednou zo základných charakteristík všeobecného modelu HERMIN je to, že ide o model malej otvorenej ekonomiky. Základný teoretický model zároveň zohľadňuje štruktúru nástrojov politiky súdržnosti. Tvorba a štruktúra tohto modelu musí spĺňať niekoľko základných predpokladov:

- Ekonomika musí byť agregovaná na malý počet sektorov, ktoré umožňujú identifikovať kľúčové štruktúrne zmeny v ekonomike počas skúmaného obdobia.
- Musí mať vyšpecifikovaný mechanizmus, pomocou ktorého je ekonomika prepojená s „vonkajším svetom“ a ktorý by mal zachytávať: medzinárodnú výmenu tovarov a služieb, inflačné prenosy, migráciu pracovnej sily, ako aj priame zahraničné investície. Vonkajšia

(alebo svetová) ekonomika je veľmi dôležitým priamym aj nepriamym faktorom, ovplyvňujúcim ekonomický rast a konvergenciu krajiny k rozvinutým krajinám EÚ.

- Produkcia jednotlivých hospodárskych sektorov zahrnutých v modeli je vyjadrená pomocou produkčných funkcií v špecifickom tvare CES (Constant Elasticity Substitution) a C-D (Cobb-Douglas).
- Tvorca a používateľ modelu musí rozpoznať možný konflikt medzi súčasným stavom ekonomiky, ktorý na základe historických dát popisuje model HERMIN, a budúcim očakávaným stavom ekonomiky približujúcej sa prostrediu dominovanému jednotným spoločným trhom EÚ

Na naplnenie týchto predpokladov sa najčastejšie používa teoretická modelová štruktúra všeobecného modelu HERMIN, ktorá sa skladá zo štyroch sektorov: sektor priemyslu, obsahujúci predovšetkým sektory ekonomiky obchodujúce so zahraničím, sektor trhových služieb, ktorý pozostáva zo sektorov ekonomiky neobchodujúcich so zahraničím a tvoriacich predovšetkým domácu ponuku, pôdohospodársky sektor a verejný sektor (nazývaný aj sektor netrhopých služieb).

Z hľadiska tvorby a použitia produkcie model pozostáva z troch blokov, a to bloku ponuky, bloku absorpcie (dopytu) a bloku rozdelenia príjmov v ekonomike. Model je skonštruovaný ako integrovaný systém rovníc s prepojeniami medzi jednotlivými blokmi a sektormi a je postavený na keynesyánskych predpokladoch a mechanizmoch, ktoré tvoria jadro modelu. V opodstatnených prípadoch obsahuje aj prvky postavené na neoklasickej ekonomickej teórii, predovšetkým v bloku ponuky. Napríklad výstupy sektora priemyslu nie sú iba jednoducho determinované dopytom, ale model zohľadňuje aj potenciálne vplyvy cenovej a nákladovej konkurencie. Odráža tak predpoklad, že firmy hľadajú miesta alebo krajiny minimalizujúce náklady produkcie. Dopyt po výrobných faktoroch v odvetviach priemyslu a trhových služieb je odvodený na základe nákladovej minimalizácie produkčnej funkcie CES (produkčná funkcia s konštantnou elasticitou substitúcie).

V bloku ponuky všeobecného modelu HERMIN je modelovaná agregátna ponuka (výstup jednotlivých sektorov), cena výstupu, index nominálnej mzdy, mzdová inflácia, konkurencie schopnosť, dopyt po práci a dopyt po investíciách. Blok taktiež obsahuje rovnice agregátnej ponuky práce, nezamestnanosti a migrácie pracovnej sily. Blok absorpcie (dopytu) obsahuje rovnice na modelovanie domácej spotreby, domáceho dopytu a saldo zahraničného obchodu. Blok rozdelenia príjmov obsahuje rovnice na výpočet cien výdavkov verejného sektora, príjmu, disponibilného príjmu domácností, verejného deficitu, verejného dlhu a saldo bežného účtu.

Aplikácia tohto modelu na regionálnej úrovni bola vyvinutá vo WARR, v Poľsku, pričom región (v prípade Poľska NUTS 2) je odhadovaný ako samostatný satelitný model s previazanosťou na národné údaje. Hlavná výhoda regionálneho ekonometrického modelu HERMIN spočíva v relatívne jednoduchej štruktúre, ktorá zároveň umožňuje analyzovanie vplyvov politiky súdržnosti na úrovni regiónov. Je úspešne aplikovateľný pre malé otvorené ekonomiky a aj v podmienkach obmedzenej údajovej základne. Jeho hlavná nevýhoda je neschopnosť zohľadniť vyššie spomínané spill-over efekty.

V slovenských podmienkach bol odhad ex-ante vplyvu odhadovaný po prvýkrát len na národnej úrovni odporúčaným modelom HERMIN (pozri Kvetan a kol., 2006). Ďalšia aplikácia tohto typu modelu bola vytvorená autormi Hudec a kol. (2009). Prvá regionálna aplikácia v slovenských podmienkach na úrovni NUTS 3 bola vytvorená v roku 2016 (pozri Radvanský a kol., 2016). Autori v tejto analýze využili schému poľského modelu, ktorý bol už aplikovaný v podmienkach geografického rozloženia NUTS 2 v Poľsku. Túto modelovú

štruktúru sme využili aj pri hodnotení efektívnosti využitia zdrojov na regionálnej úrovni v tomto článku.

Objemy finančných prostriedkov implementované v jednotlivých regiónoch boli identifikované na základe spracovania údajov o realizovaných projektoch z databázy ITMS. V prípade projektov realizovaných v jednom NUTS 3 regióne išlo o relatívne jednoduchú úlohu, v prípade nadregionálnych projektov bolo potrebné realizovať dezagregáciu pridelených zdrojov medzi regióny na území ktorých sa projekt realizoval. Pri kvantifikácii čerpania ŠF a KF v rámci nadregionálnych projektov boli zohľadnené špecifiká operačných programov a podporovaných projektov, čo umožnilo presnejšiu dezagregáciu údajov o čerpaní na regionálnu úroveň.

2. HLAVNÉ ZISTENIA

Efektívnosť geografického rozloženia finančných prostriedkov ŠF a KF sme posudzovali z viacerých hľadísk. V prvom rade geografické rozdelenie čerpania prostriedkov ŠF a KF na úrovni regiónov NUTS 3 do značnej miery odzrkadľuje ekonomickú výkonnosť jednotlivých regiónov (snaha o regionálnu konvergenciu) a charakter projektov realizovaných na ich území (riešenie problémových oblastí, napr. investičné dlhy).

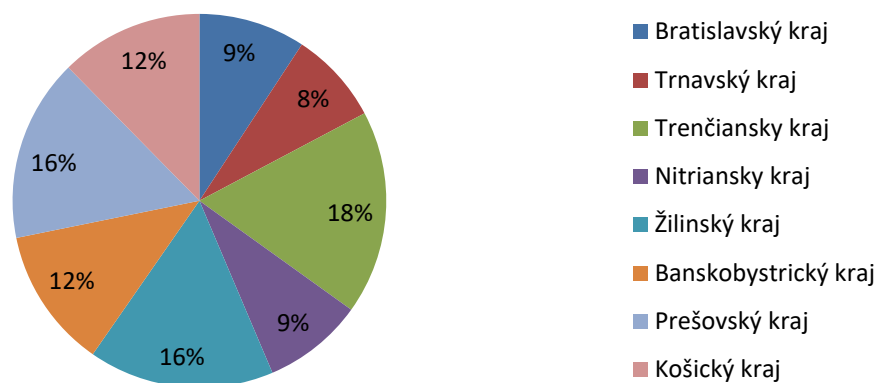
S ohľadom na limitovanú oprávnenosť čerpania finančných prostriedkov a relatívne nízky podiel čerpania v Bratislavskom kraji (BA), bolo v tomto kraji ku koncu roka 2014 vyčerpaných približne 9 % celkového objemu využiteľných zdrojov. Vzhľadom na zvýšené tempo čerpania (investície do dopravnej infraštruktúry a modernizácie hromadnej dopravy) v priebehu roku 2014 odsunul Bratislavský kraj v objeme čerpania na posledné miesto Trnavský kraj (TT). Z dôvodu poľnohospodárskeho zamerania produkcie² na väčšine územia Nitrianskeho kraja (NR) čerpal tento región podobný objem finančných prostriedkov zo ŠF a KF ako Bratislavský kraj. Vyššia koncentrácia veľkých infraštruktúrnych projektov realizovaných prevažne v severo-južnej osi Považia mala za následok, že takmer 18 % finančných prostriedkov zo ŠF a KF bolo do konca roka 2014 vyčerpaných na území Trenčianskeho kraja (TN). V Trenčianskom kraji výdavky na infraštruktúru tvorili viac ako 76 % celkových výdavkov zo ŠF a KF smerujúcich do tohto kraja. Relatívne vysokú koncentráciu výdavkov na infraštruktúru sme zaznamenali aj v Žilinskom (ZA) a Prešovskom (PO) kraji, v ktorých tvorili 62 %, resp. 60 % celkových výdavkov. Ďalšiu dvojicu krajov s podobnou úrovňou čerpania tvoria Banskobystrický (BB) a Košický (KE) kraj, v ktorých bolo vyčerpaných približne 12 % celkových výdavkov zo ŠF a KF. Podobne ako v prípade severnej dvojice krajov východnej časti Slovenska, aj pri týchto krajoch tvorili výdavky na infraštruktúru približne polovicu celkového čerpania (52 % BB, resp. 47,7 % KE). Z pohľadu objemov čerpaných prostriedkov a ich štruktúry existujú značné rozdiely v geografickom rozdelení čerpania finančných prostriedkov na úrovni krajov NUTS 3, ktoré boli determinované sledovanými cieľmi.

Ďalším dôležitým aspektom pri skúmaní efektívnosti geografickej alokácie je pohľad na schopnosť jednotlivých regiónov čerpať dostupné finančné prostriedky zo ŠF a KF – absorpčná kapacita regiónov. Pri detailnom pohľade na priebeh čerpania finančných prostriedkov fondov EÚ počas programového obdobia 2007 – 2013 je zjavné, že došlo k oneskoreniu začatia reálneho čerpania. V prvých rokoch programového obdobia, t.j. 2007 a 2008, bolo čerpanie iba marginálne vzhľadom na rozsah dostupných zdrojov, pričom

² Tento typ produkcie je finančne podporovaný z iných fondov EÚ.

podstatnú časť výdavkov tvorili prostriedky technickej pomoci využívané subjektmi zapojenými do implementácie politiky súdržnosti.

Graf 1: Podiely jednotlivých krajov na čerpaní ŠF a KF



Zdroj: ITMS

Až rok 2009 je možné považovať za rok začiatku reálneho čerpania finančných prostriedkov zo ŠF a KF, pričom prevažná časť prostriedkov čerpaných v tomto roku smerovala na naštartovanie infraštruktúrnych projektov v Trenčianskom, Žilinskom a Banskobystrickom kraji. Podobný stav sme zaznamenali aj v roku 2010, v ktorom sa začali realizovať vo väčšej miere infraštruktúrne projekty aj v Prešovskom kraji. Od roku 2011 došlo k výraznému zintenzívneniu tempa čerpania finančných prostriedkov fondov EÚ, pričom prvý krát v tomto roku prekročil vo všetkých krajoch objem čerpaných prostriedkov sumu 100 mil. EUR. Do roku 2012 bol viditeľný trend rastu intenzity čerpania vo všetkých regiónoch Slovenska. V roku 2013 sa tento trend v prípade štyroch krajov zastavil (TT, NR, ZA a BB) a došlo k miernemu poklesu čerpaných finančných prostriedkov v porovnaní s rokom 2012. To by mohlo indikovať, že aktuálna úroveň čerpania finančných prostriedkov zo ŠF a KF na úrovni niektorých krajov je blízko k ich absorpčnej kapacite.

Tabuľka 1: Podiel čerpaných prostriedkov zo zdroja EÚ na HDP kraja v %

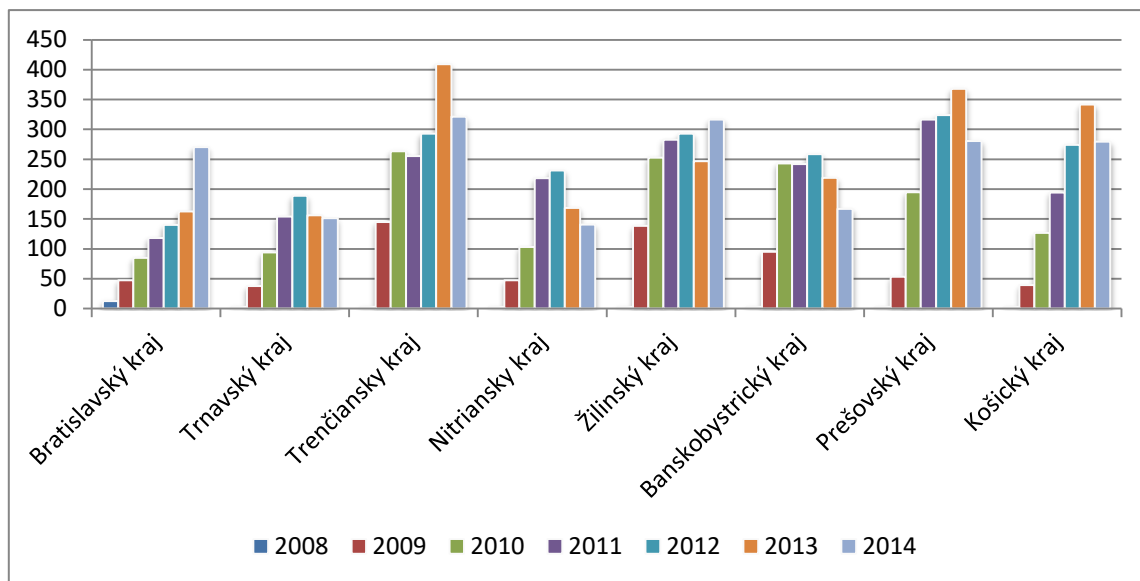
	2007	2008	2009	2010	2011	2012	2013	2014	2015
BA	0,0	0,1	0,2	0,4	0,5	0,6	0,7	1,2	1,2
TT	0,0	0,0	0,4	0,9	1,5	1,8	1,6	1,5	2,6
TN	0,0	0,0	2,0	3,4	3,0	3,3	5,1	4,2	7,3
NR	0,0	0,0	0,5	1,2	1,9	2,1	1,6	1,4	2,8
ZA	0,0	0,0	1,6	2,9	3,0	3,0	2,8	3,6	5,8
BB	0,0	0,0	1,4	3,3	2,9	3,2	2,9	2,2	5,5
PO	0,0	0,0	0,8	2,6	4,1	4,2	5,0	3,9	7,0
KE	0,0	0,0	0,4	1,3	2,0	2,7	3,5	2,2	4,1
SR	0,0	0,0	0,8	1,6	2,0	2,2	2,4	2,2	3,8

Zdroj: výpočty autorov

Tento fakt je podporený aj relatívne vyššou efektívnosťou čerpaných finančných prostriedkov zo ŠF a KF v prípade Bratislavského kraja, v ktorom tieto prostriedky tvorili menej ako 1 %

regionálneho HDP. V priebehu roku 2014 došlo k viacerým komplikáciám v čerpaní prostriedkov ŠF a KF, ktoré sa napriek očakávanej akcelerácii prejavili naopak medziročným poklesom tempa čerpania vo všetkých regiónoch s výnimkou BA a ZA. Jedným zo základných problémov programového obdobia 2007 – 2013 bolo oneskorenie čerpania finančných prostriedkov v rokoch 2007 a 2008, ktoré bolo len veľmi komplikovane korigované v poslednom roku s možnosťou čerpania zdrojov. Medzi ďalšie problémy limitujúce úroveň čerpania v krajoch patrila nedostatočná pripravenosť projektových výziev na začiatku programového obdobia, komplexnosť problematiky čerpania, relatívne vysoká administratívna náročnosť a ťažko splniteľné podmienky niektorých výziev, a to predovšetkým vo väzbe na malé a stredné podniky.

Graf 2: Čerpanie finančných prostriedkov podľa regiónov NUTS 3 vrátane kofinancovania, mil. EUR



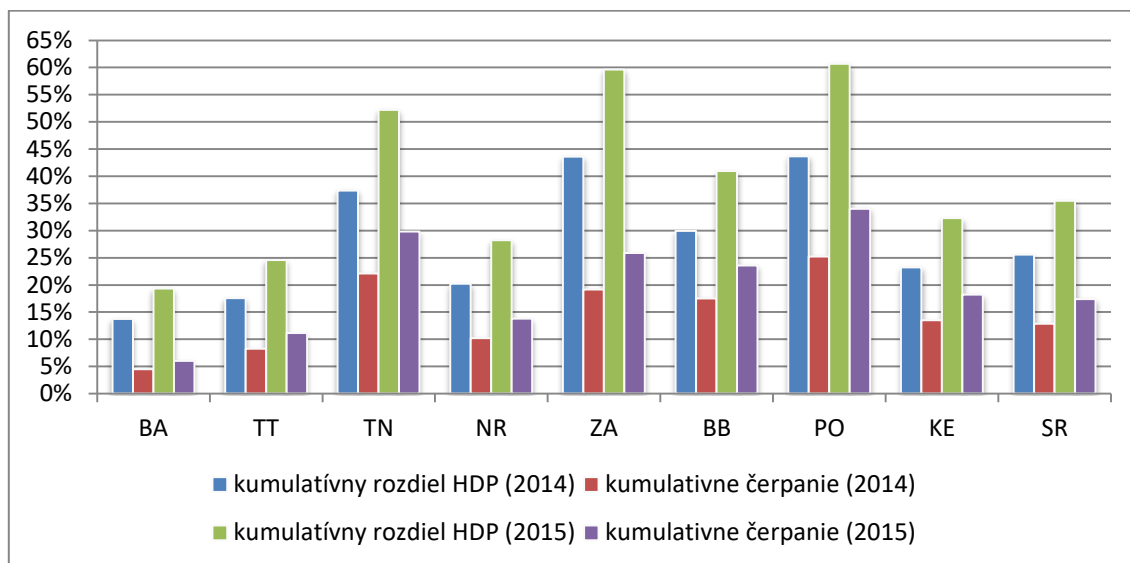
Zdroj: autori na základe dát z ITMS

Dôležitú informáciu o efektívnosti čerpania nám poskytuje pohľad na kumulatívnu hodnotu nárastu HDP v dôsledku implementácie ŠF a KF na úrovni jednotlivých krajov. Do konca roku 2014 bol najvyšší podiel dodatočného kumulatívneho HDP v porovnaní s úrovňou HDP v roku 2007 zaznamenaný v prípade Žilinského (38 %), Prešovského (35 %) a Trenčianskeho kraja (25 %). Najnižší podiel bol v roku 2014 zaznamenaný v Bratislavskom kraji, v ktorom dosiahol približne 12 %. Obdobný vývoj je možné vidieť z grafu 3 aj v prípade dopadov v roku 2015, pričom v Prešovskom, Trenčianskom a Žilinskom kraji je možné očakávať vytvorenie dodatočného HDP úhrne predstavujúceho viac ako polovicu úrovne HDP z prvého roku programového obdobia. Najnižší objem dodatočného HDP vytvoreného implementáciou fondov EÚ v roku 2015, na úrovni 19 % HDP roku 2007, sa realizoval v Bratislavskom kraji. Je však nutné poznamenať, že objem prostriedkov čerpaných v tomto regióne predstavuje len zhruba 6 % HDP roku 2007 a teda efektívnosť použitia prostriedkov ŠF a KF v tomto regióne je relatívne vysoká.

Na základe hodnôt CSF multiplikátora sa ako najefektívnejší javí Bratislavský kraj, ktorý predstavuje ekonomicky najsilnejší región v SR a jeho absorpčný potenciál aj napriek značnej akcelerácii v posledných rokoch programového obdobia vysoko prevyšuje objem alokovaných prostriedkov. Rovnako aj charakter projektov realizovaných v rámci BA zahŕňalo predovšetkým aktivity s vyššou pridanou hodnotou, ktoré generujú vyššie multiplikačné efekty. Multiplikačné efekty sú v tomto kraji relatívne vyššie aj z dôvodu, že

veľká časť čerpaných zdrojov bola použitá na mzdy v pracovne náročnejších sektoroch služieb. Tieto mzdy následne relatívne rýchlo generujú cez rastúci dopyt domácností dodatočné nepriame efekty v ekonomike regiónu. V tomto kraji dosiahol na konci roku 2014 CSF multiplikátor relatívne vysokú hodnotu 3,1, avšak podiel čerpania na HDP až v roku 2014 prekročil 1% HDP, čo sa prejavilo na relatívne nižších efektoch v absolútnom vyjadrení. Skupinu krajov s vyššou mierou efektívnosti predstavujú Žilinský, Trnavský a Nitriansky kraj, ktoré profitovali zo svojej ekonomickej sily a relatívne vyššieho podielu výdavkov na podporu priemyslu a služieb.

Graf 3: Porovnanie kumulatívneho čerpania a dodatočného HDP v rokoch 2014 a 2015 podľa regiónov NUTS 3 a SR, podiel na HDP roku 2007



Zdroj: výpočty autorov, ITMS

Najnižšiu efektívnosť meranú multiplikátorom CSF z ekonomicky silnejších krajov dosiahol Trenčiansky kraj, čo bolo spôsobené vysokým podielom výdavkov na infraštruktúru, pri ktorých sa multiplikačné efekty dostavia až s väčším časovým posunom. Investície do infraštruktúry so sebou prinášajú aj nižšie nepriame efekty na hospodárstvo regiónu vzhľadom na relatívne nižší objem mzdových prostriedkov. Podobnú efektívnosť dosiahol aj ekonomicky slabší Košický kraj, v prípade ktorého efektívnosť do značnej miery potiahli výdavky na výskum a vývoj a na pomoc priemyslu a službám.

Tabuľka 2: Kumulatívny multiplikátor CSF podľa regiónov NUTS 3

	2008	2009	2010	2011	2012	2013	2014	2015
Bratislavský kraj	2,2	2,5	2,7	2,6	2,7	3,0	3,1	3,2
Trnavský kraj	0,8	1,0	1,3	1,5	1,8	2,0	2,1	2,2
Trenčiansky kraj	0,9	1,2	1,3	1,4	1,5	1,6	1,7	1,8
Nitriansky kraj	0,8	1,0	1,1	1,4	1,6	1,8	2,0	2,0
Žilinský kraj	1,2	1,6	1,8	1,9	2,0	2,2	2,3	2,3
Banskobystrický kraj	0,8	1,0	1,1	1,2	1,4	1,6	1,7	1,7
Prešovský kraj	0,9	1,1	1,2	1,3	1,5	1,6	1,7	1,8
Košický kraj	0,9	1,0	1,2	1,3	1,6	1,6	1,7	1,8
SR	1,4	1,3	1,4	1,5	1,7	1,8	2,0	2,0

Zdroj: výpočty autorov

Poslednú skupinu s podobnými hodnotami CSF multiplikátora tvoria ekonomicky najslabšie kraje Banskobystrický a Prešovský, v ktorých bola značná časť prostriedkov smerovaná do

budovania a modernizácie infraštruktúry, čím prispeli k znižovaniu ich investičného dlhu. Za pozitívum je možné považovať fakt, že hodnota multiplikátora sa v prípade všetkých regiónov postupne zvyšovala a v prípade žiadneho regiónu nebola v roku 2014 nižšia ako 1. Ako vysoko efektívne môžeme označiť regióny, v ktorých dosiahla úroveň CSF multiplikátora hodnotu vyššiu ako 2. Táto hranica predstavovala podľa Bradley a Untiedt (2009) úroveň CSF multiplikátora, ktorú dosahovali stredne efektívne krajiny v programovom období 2000 – 2006. Hlavným zdrojom rozdielov regionálneho multiplikátora je štruktúra implementovaných projektov, pričom vyššia efektívnosť bola zaznamenaná v regiónoch, v ktorých sa realizovali projekty s relatívne vyššou pracovnou náročnosťou a mierou pridanej hodnoty.

3. DISKUSIA A ZÁVERY

Dosiahnuté výsledky naznačujú, že geografické rozdelenie finančných zdrojov ŠF a KF v programovom období 2007-2013 sledovalo ciele definované politikou súdržnosti. Ako najefektívnejší v čerpaní zdrojov politiky súdržnosti je možné vyhodnotiť Bratislavský kraj, čo však bolo determinované v prevažnej miere sektorovou štruktúrou prostriedkov čerpaných v tomto kraji, ktorá bola relatívne významne zameraná na podporu vedy a výskumu, služieb a ľudských zdrojov v inštitúciách zodpovedných za riadenie a implementáciu prostriedkov fondov EÚ. Až v roku 2014 sa do štruktúry investícií v tomto kraji pridali aj výdavky na podporu priemyselnej produkcie.

Hodnoty CSF multiplikátora v jednotlivých regiónoch indikujú, že výdavky na podporu priemyslu a služieb prinášajú intenzívnejšie okamžité efekty na rast HDP, zatiaľ čo investície do infraštruktúry sa prejavujú až s relatívne značným oneskorením. Hoci sú investície do infraštruktúry z krátkodobého až strednodobého hľadiska menej efektívne, predstavujú nevyhnutný predpoklad pre budúci ďalší ekonomický rast jednotlivých regiónov. Z hľadiska regionálnej alokácie je teda potrebné brať do úvahy konflikt medzi rastovou a kohéznou alokáciou ŠF a KF v programovom období 2007-2013. Geografická efektívnosť rozdelenia zdrojov meraná CSF multiplikátorom indikuje snahu o plnenie cieľov NSRR, a finančné prostriedky boli do značnej miery alokované geograficky efektívne.

Výsledky efektívnosti merané hodnotou CSF multiplikátora ukázali, že v žiadnom z krajov neboli prostriedky fondov EÚ v programovom období 2007-2013 využité neefektívne, nakoľko vo všetkých regiónoch dosiahla hodnota CSF multiplikátora hodnotu vyššiu ako 1. Najvyššiu efektívnosť sme zaznamenali v prípade Bratislavského a Žilinského kraja v prípade ktorých sa pozitívne prejavili najmä investície do oblastí vedy a výskumu. Za relatívne vysoko efektívne možno považovať kraje v prípade ktorých dosiahla úroveň CSF multiplikátora hodnotu vyššiu ako 2 (BA, TT, NR a ZA).

Je zrejmé, že ciele politiky súdržnosti a podpory ekonomického rastu na národnej a regionálnej úrovni môžu byť protichodné z hľadiska alokácií ako aj predpokladanej efektívnosti čerpania. Je preto dôležité dať do súladu cieľ politiky súdržnosti a jeho vplyvu na ekonomický vývoj na národnej a regionálnej úrovni (národná vs. regionálna konvergencia) v rámci tvorby ucelenej rozvojovej politiky.

POUŽITÁ LITERATÚRA

1. BARRY, F., BRADLEY, J., KEJAK, M., Vavra, D. 2003. The Czech economic transition: exploring options using a macrosectoral model, *The Economics of Transition*, vol. 11(3), s. 539-567.

2. BRADLEY, J. 2000. Policy Design and Evaluation: EU Structural Funds and Cohesion in the European Periphery, In: Empirical Models and Policy-Making: Interaction and Institutions, F. den Butter and M. Morgan (eds.), London: Routledge.
3. BRADLEY, J., GÁCS, J. KANGUR, A., LUBENETS, N. 2005. HERMIN: A Macro Model Framework for the Study of Cohesion and Transition, In: Bradley J, Petrakos G. and Traistaru I. (eds) Integration, Growth and Cohesion in an Enlarged European Union, Springer, Berlin, s. 207- 243.
4. BRADLEY, J., HERCE, J. A., MODESTO, L. 1995. Modelling in the EU periphery : The HERMIN project, Economic Modelling, Elsevier, vol. 12(3), pages 219-220, July.
5. D'ALCANTARA, G., ITALIANER, A. 1982. A European Project for a Multinational Macrosectoral Model, Document MS 11, DG XII, European Commission, 1982, Brussels.
6. HUDEC, O., URBANČÍKOVÁ, N., DŽUPKA, P., ŠEBOVÁ, M., KLIMOVSKÝ, D., SUHÁNYI, L., ŽELINSKÝ, T. 2009. Podoby regionálneho a miestneho rozvoja, Ekonomická fakulta, TU Košice. ISBN 978-80-553-0117-4
7. KVETAN, V., MLÝNEK, M., RADVANSKÝ, M. 2006. Model HERMIN, Ekonomický časopis/Journal of Economics, vol. 54 (10), pages 994-1008.
8. RADVANSKÝ, M. - FRANK, K. 2010. Modelling the impact of postponed implementation of EU structural funds - the case Slovakia, In: EcoMod2010 : international conference on economic modeling [online]. - Istanbul : Istanbul Bilgi University, 2010, 11 s. Dostupné na: <http://www.ecomod.org/files/papers/1371.pdf>
9. RADVANSKÝ, M., FRANK, K., LICHNER, I., MIKLOŠOVIČ, T. 2016. Impact of cohesion policy on regional development of Slovakia : ex-post assessment of national strategic reference framework 2007-2013. Bratislava : Institute of Economic Research SAS, 2016. 111 s. ISBN 978-80-7144-257-8. Dostupné na: <<http://ekonom.savba.sk/sk/publikacie/-p334>>.
10. VARGA, J., IN'T VELD, J. 2009. A Model-based Analysis of the Impact of Cohesion Policy Expenditure 2000-06: Simulations with the QUEST III endogenous R&D model, European Commission, Brussels.
11. ZALESKI, J., TOMASZEWSKI, P., WOJTASIAK, A., BRADLEY, J. 2004. Revision of the Polish HERMIN model using the new national accounts: 1995-2002, Wroclaw Regional Development Agency Working paper, Wroclaw.
12. ZALESKI, J. 2009. Regionalization of HERMIN macro-economic modeling framework in Poland, Wroclaw Regional Development Agency Working paper, Wroclaw.

Kontaktné údaje

Ing. Marek Radvanský, PhD.

Mgr. Tomáš Miklošovič, PhD.

Ing. Ivan Lichner, PhD.

Institute of Economic Research, Slovak Academy of Sciences

Šancová 56, 811 05 Bratislava

Slovak Republic

email: marek.radvansky@savba.sk, tomas.miklosovic@savba.sk, ivan.lichner@savba.sk

ÚVOD DO EKONOMETRICKEJ ANALÝZY SO SYSTÉMOM R¹

INTRODUCTION TO ECONOMETRIC ANALYSIS WITH R SYSTEM

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 základy ekonometrickej analýzy s programovacím systémom R, ktorý je voľne dostupný a predstavuje kvalitnú alternatívu komerčným špecializovaným softvérovým produktom.

Kľúčové slová: ekonometrická analýza, programovací systém R, RStudio

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 the basics of econometric analysis with R programming system, which is freely available and it is a quality alternative to commercial software product.

Keywords: econometric analysis, R programming system, RStudio

1 ÚVOD

Postup ekonometrickej analýzy si vysvetlíme na dvoch príkladoch lineárnych regresných modelov, pričom využijeme údaje o výdavkoch domácností na spotrebu a disponibilnom príjme, ktoré uvádza Hatrák (2003). Údaje k prvému príkladu sa dajú stiahnuť napríklad na internetovej stránke <http://fhi.sk/files/katedry/kove/ucebnice/EMEG/Kap02-pr5.xls>.

Okrem tohto príkladu využívajúceho prierezné dáta, na ktorom ukážeme testovanie heteroskedasticity, prezentujeme aj príklad s časovými radmi a viacerými vysvetľujúcimi premennými o spotrebe cigariet v závislosti od reálneho príjmu na osobu a reálnej ceny cigariet, ktorý uvádza Ramanathan (1998), na ktorom ukážeme testovanie autokorelácie. Údaje k druhému príkladu sa nachádzajú v hárku Kap7Tab31 v súbore Ekonometria1.xls na internetovej stránke <http://fhi.sk/sk/katedry-fakulty/kove/ucebnice/ekon1>.

Najskôr uložíme súbor Kap02-pr5.xls v Exceli vo formáte typu .csv. Skôr ako urobíme to isté s hárkom Kap7Tab31 súboru Ekonometria1.xls, dopíšeme nad roky do prvej bunky názov stĺpca *ROK*. Následne môžeme súbory s dátami, ktoré sú podľa slovenských pravidiel (oddeľovač bodkočiarka a pre desatiny čiarka) importovať ako Lukáčik et al (2016) príkazmi:

```
> hatrak <- read.csv(file = "Kap02-pr5.csv", sep=";")
> ramanathan <- read.csv(file = "Ekonometria1.csv", sep=";", dec = ",")
> attach(hatrak)
> attach(ramanathan)
```

¹ Príspevok vznikol s podporou projektov VEGA 1/0444/15 "Ekonometrická analýza produkčných možností ekonomiky a trhu práce na Slovensku" a VEGA 1/0028/14 "Regionálne modelovanie ekonomického rastu krajín EÚ s dôrazom na modely priestorovej ekonometrie".

2 ODHAD PARAMETROV LINEÁRNYCH MODELOV

Pri odhade parametrov lineárneho regresného modelu pomocou metódy najmenších štvorcov (MNS) môžeme využiť silnú podporu maticového zápisu systému R alebo jeho funkcie.

Predpokladajme pre výdavky domácností na spotrebu model v tvare:

$$\text{spotreba}_i = \beta_0 + \beta_1 \text{prijem}_i + u_i \quad (1)$$

a pre spotrebu cigariet logaritmický model v tvare:

$$\ln Q_t = \beta_0 + \beta_1 \ln Y_t + \beta_2 \ln P_t + \beta_3 D82_t + u_t \quad (2)$$

Pre vektor odhadnutých parametrov $\hat{\beta}$ platí podľa MNS známy vzťah:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3)$$

kde \mathbf{X} je matica vysvetľujúcich premenných, ktorej prvý stĺpec v prvom príklade je vektor jednotiek, druhý stĺpec je vektor **prijem** a \mathbf{y} je vektor **spotreba**. V druhom prípade je opäť prvý stĺpec vektor jednotiek, druhý stĺpec je vektor logaritmov \mathbf{Y} , tretí stĺpec je vektor logaritmov \mathbf{P} , štvrtý stĺpec je vektor umelej premennej **D82** a \mathbf{y} je vektor logaritmov \mathbf{Q} .

Vytvoríme maticu vysvetľujúcich premenných \mathbf{Xh} pre model (1) a maticovými operátormi (*solve()* pre inverziu matice, *t()* pre transpozíciu matice a *%*%* pre súčin dvoch matic) vypočítame vektor $\hat{\beta}$ (označíme ho **beta_vyr1**). Následne model odhadneme funkciou *lm* a jeho odhad zapíšeme do objektu *lin_mod1*. Porovnaním vidíme, že výsledky sa zhodujú.

Vytvoríme maticu vysvetľujúcich premenných \mathbf{Xr} pre model (2) a rovnakými maticovými operátormi vypočítame vektor $\hat{\beta}$ (označíme ho **beta_vyr2**). Následne model odhadneme funkciou *lm* a jeho odhad zapíšeme do objektu *lin_mod2*. Porovnaním opäť vidíme, že výsledky sa zhodujú.

Postup odhadu parametrov modelu (1) maticovo a zabudovanou funkciou *lm*:

```
> Xh <- cbind(1, prijem) # vytvoríme maticu vysvetľujúcich premenných Xh
> beta_vyr1 <- solve(t(Xh) %*% Xh) %*% t(Xh) %*% spotreba # odhad vektora beta_vyr1
> lm(spotreba ~ prijem) # odhad parametrov lineárneho modelu funkciou
> lin_mod1 <- lm(spotreba ~ prijem) # odhad zapíšeme do objektu lin_mod1
> summary(lin_mod1) # komplexnejší výpis informácií o parametroch z odhadu
```

Postup odhadu parametrov modelu (2) maticovo a zabudovanou funkciou *lm*:

```
> Xr <- cbind(1, log(Y), log(P), D82) # vytvoríme maticu vysvetľujúcich premenných Xr
> beta_vyr2 <- solve(t(Xr) %*% Xr) %*% t(Xr) %*% log(Q) # odhad vektora beta_vyr2
> lm(log(Q) ~ log(Y) + log(P) + D82) # odhad parametrov lineárneho modelu funkciou
> lin_mod2 <- lm(log(Q) ~ log(Y) + log(P) + D82) # odhad zapíšeme do objektu lin_mod2
> summary(lin_mod2) # komplexnejší výpis informácií o parametroch z odhadu
```

Odhadnutý model (1) pre výdavky domácností na spotrebu má tvar:

$$\begin{array}{llll} \text{spotreba}_i = -632,93 + 0,7267 \text{prijem}_i & n = 50 & & \\ (922,08) & (0,046) & \hat{\sigma} = 1865 & F = 245,5 \\ [-0,686] & [15,67] & R^2 = 0,8365 & R_{adj}^2 = 0,8330 \end{array}$$

Pod hodnotami odhadnutých parametrov sú zapísané odhady štandardných odchýlok parametrov (v okrúhlych zátvorkách) a t-štatistík (v hranatých zátvorkách).

Odhadnutý model (2) pre spotrebu cigariet má tvar:

$$\ln Q_t = -4,636 + 0,684 \ln Y_t - 0,322 \ln P_t - 0,125 D82_t \quad n = 29$$

(0,576)	(0,075)	(0,089)	(0,028)	$\hat{\sigma} = 0,0472$	$F = 41$
[-8,050]	[9,085]	[-3,630]	[-4,404]	$R^2 = 0,8311$	$R^2_{adj} = 0,8108$

```

Console Z:/FHI_EU/Documents/Prispevky/Praha2016/
> Xh <- cbind(1, prijem)
> beta_vyr1 <- solve(t(Xh) %*% Xh) %*% t(Xh) %*% spotreba
> beta_vyr1
      [,1]
-632.9266479
prijem 0.7267341
> lm(spotreba ~ prijem)
Call:
lm(formula = spotreba ~ prijem)

Coefficients:
(Intercept)      prijem
-632.9266      0.7267

> lin_mod1 <- lm(spotreba ~ prijem)
> summary(lin_mod1)

Call:
lm(formula = spotreba ~ prijem)

Residuals:
    Min       1Q   Median       3Q      Max
-3781.1 -1000.1   54.4  1287.4  4630.9

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -632.92665    922.07874   -0.686   0.496
prijem       0.72673     0.04638    15.668 <2e-16 ***
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1865 on 48 degrees of freedom
Multiple R-squared:  0.8365, Adjusted R-squared:  0.833
F-statistic: 245.5 on 1 and 48 DF, p-value: < 2.2e-16

Console Z:/FHI_EU/Documents/Prispevky/Praha2016/
> Xr <- cbind(1, log(Y), log(P), D82)
> beta_vyr2 <- solve(t(Xr) %*% Xr) %*% t(Xr) %*% log(Q)
> beta_vyr2
      [,1]
-4.6360081
 0.6841978
-0.3220709
D82 -0.1245595
> lm(log(Q) ~ log(Y) + log(P) + D82)
Call:
lm(formula = log(Q) ~ log(Y) + log(P) + D82)

Coefficients:
(Intercept)      log(Y)      log(P)      D82
-4.6360      0.6842     -0.3221     -0.1246

> lin_mod2 <- lm(log(Q) ~ log(Y) + log(P) + D82)
> summary(lin_mod2)

Call:
lm(formula = log(Q) ~ log(Y) + log(P) + D82)

Residuals:
    Min       1Q   Median       3Q      Max
-0.12045 -0.02273 -0.01360  0.01795  0.12167

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.63601    0.57589   -8.050 2.10e-08 ***
log(Y)       0.68420    0.07531   9.085 2.14e-09 ***
log(P)      -0.32207    0.08872   -3.630 0.001272 **
D82         -0.12456    0.02828   -4.404 0.000175 ***
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04719 on 25 degrees of freedom
Multiple R-squared:  0.8311, Adjusted R-squared:  0.8108
F-statistic: 41 on 3 and 25 DF, p-value: 8.361e-10

```

Obrázok 1: Odhad parametrov oboch lineárnych modelov maticovo a funkciu

3 ĎALŠIE VÝSLEDKY ODHADU VYPOČÍTANÉ FUNKCIOU LM

Pri odhade modelu pomocou funkcie sa vypočíta až 12 rôznych zložiek objektu. Pozrieť si ich môžeme v okne *Environment* programu RStudio alebo vypísať príkazom *names(lin_mod1)*. Vypočítané sú: *coefficients*, *residuals*, *effects*, *rank*, *fitted.values*, *assign*, *qr*, *df.residual*, *xlevels*, *call*, *terms* a *model*.

Na obr. 1 vidíme, aké informácie získame pomocou príkazu *summary(lin_mod1)*. Ak by sme priradili *summary()* do objektu s názvom napríklad *out_lin_mod1* príkazom:

```
> out_lin_mod1 <- summary(lin_mod1)
```

a pre tento objekt si pozreli jeho zložky príkazom *names(out_lin_mod1)*, vidíme, že priradené sú: *call*, *terms*, *residuals*, *coefficients*, *aliased*, *sigma*, *df*, *r.squared*, *adj.r.squared*, *fstatistic* a *cov.unscaled*. Niektoré zložky sú totožné so zložkami objektu *model*. Výhodou priradenia k objektu je, že máme k nim priamy prístup cez názov objektu, za ktorý dáme znak dolára a zaň zapíšeme názov zložky. Napríklad *F* štatistika hypotézy, či sa všetky parametre modelu okrem konštanty súčasne rovnajú 0, vypísaná na konci *summary* s *p-value* hodnotou, sa získa:

```
> out_lin_mod1$fstatistic
```

a okrem jej presnejšej hodnoty 245,4968 sa zobrazia aj stupne voľnosti čitateľa a menovateľa.

Tabuľka s odhadnutými parametrami, odhadmi ich štandardných hodnôt, *t* štatistikami a pravdepodobnostnými hodnotami je v zložke *coefficients*, ktorý je maticou. To znamená, že na jej prvky sa odvoláme ich pozíciou, napr. štandardná odchýlka parametra sklonu k spotrebe je *out_lin_mod1\$coefficients[2,2]*.

Podobne $(\mathbf{X}^T \mathbf{X})^{-1}$ je v zložke *cov.unscaled*, ktorá je maticou a na jej prvky sa odvoláme ich pozíciou. Ak by sme jej pomocou chceli získať rozptyl parametra sklonu k spotrebe, musíme najskôr vypočítať variančno-kovariančnú maticu $\mathbf{V} = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$ príkazom:

```
> V <- out_lin_mod1$cov.unscaled*out_lin_mod1$sigma^2
```

a následne sa odvoláme na jeho pozíciu *V[2,2]*.

Pre odhadnutý model existuje okrem *summary()* niekoľko ďalších tzv. generických funkcií:

- *coef()* (alebo *coefficients()*) vyberie odhadnuté parametre,
- *resid()* (alebo *residuals()*) vyberie reziduály,
- *fitted()* (alebo *fitted.values()*) vyberie vyrovnané hodnoty,
- *confint()* konfidenčné intervaly odhadnutých parametrov,
- *predict()* prognózované hodnoty pre nové údaje,
- *vcov()* odhadnutá variančno-kovariančná matica (netreba rátať \mathbf{V} cez *cov.unscaled*),
- *anova()* analýza rozptylu,
- *plot()* diagnostické grafy,
- *deviance()* súčet štvorcov reziduálov,
- *logLik()* logaritmus vierohodnostnej funkcie,
- *AIC()* prípadne *BIC()* Akaikeho príp. Schwarzovo informačné kritérium.

Potom napr. bodový odhad (druhá časť výpisu za *call* pri zadaní odhadu modelu funkciou) sa dá získať príkazom:

```
> coef(lin_mod1)
```

a intervalový odhad na 5 % hladine významnosti zasa príkazom

```
> confint(lin_mod1,level = 0.95)
```

Pri prognózovaní je vhodné vytvoriť nový dátový rámec, aby sme neprepísali skutočné hodnoty vysvetľujúcich premenných. Keďže najväčšia hodnota príjmu v údajoch je 29721, zadáme nové hodnoty ako sekvenciu od 30000 po 54000 zo skokom 1000 príkazom:

```
> new_x <- data.frame(prijem=seq(30000,54000,1000))
```

a následne prognózuje spotrebu pre tieto hodnoty príjmu vrátane štandardných odchýlok. Argumentom *interval = "confidence"* vypočítame intervalové prognózy pre stredné hodnoty vysvetľujúcej premennej (ležiace na regresnej priamke) a argumentom *interval = "prediction"* vypočítame intervalové prognózy pre individuálne hodnoty vysvetľujúcej premennej:

```
> predict(lm(spotreba ~ prijem), new_x, se.fit = TRUE, interval = "prediction")
```

Pri prognóze jednej hodnoty (napr. pre *prijem = 20000*) môžeme oba kroky spojiť do jedného:

```
> predict(lm(spotreba ~ prijem), data.frame(prijem=20000), se.fit=TRUE, interval = "prediction")
```

Pri práci s časovými radmi môže byť výhodné premennú zaviesť ako objekt typu *ts*, lebo vtedy môžeme využiť diferencie a oneskorenia premenných. V tomto type objektu sa definuje prvé a posledné obdobie a frekvencia dát (ročné = 1, štvrťročné = 4, mesačné = 12). Naše dáta:

```
> Q <- ts(data = ramanathan$Q, start=1960, end=1988, frequency = 1)
```

```
> Y <- ts(data = ramanathan$Y, start=1960, end=1988, frequency = 1)
```

```
> P <- ts(data = ramanathan$P, start=1960, end=1988, frequency = 1)
```

```
> D82 <- ts(data = ramanathan$D82, start=1960, end=1988, frequency = 1)
```


4 GRAFY ZÁVISLOSTI PREMENNÝCH A VÝSLEDKOV ODHADU

Okrem odhadov nám programovací systém R môže posloužiť aj na vizualizáciu dát pomocou grafov. Základným príkazom zobrazenia vzájomnej závislosti dvoch premenných je *plot*. Ako prvý argument zadávame premennú na osi *x* a následne premennú na osi *y*. Ďalšie argumenty špecifikujú tvar znaku označujúceho hodnoty (*pch*), ich farbu (*col*), nadpis grafu (*main*), mierky osí alebo legendu.

```
> plot(prijem, spotreba, pch=16, col="blue", main="Výdavky domácností") # graf závislosti
```

Ak sme odhadli parametre modelu pomocou funkcie *lm*, môžeme zobrazit' na grafe aj regresnú priamku príkazom *abline* s prvým argumentom názvom modelu. Ďalšie argumenty špecifikujú vlastnosti regresnej priamky, napríklad jej hrúbku (*lwd*).

```
> abline(lin_mod1, lwd=2) # pridá do grafu regresnú priamku
```

Ak sme odhadli model pomocou funkcie *lm*, môžeme zobrazit' na tomto grafe aj reziduály.

Na zobrazenie reziduálov na tomto bodovom grafe si potrebujeme vytvorit' premenné, do ktorých si zapamätáme vektor hodnôt reziduálov a vektor vyrovnaných hodnôt.

```
> res <- signif(residuals(lin_mod1), 5) # zapamätáme si zaokrúhlené reziduály
```

```
> pred <- predict(lin_mod1) # zapamätáme si vyrovnané hodnoty spotreby
```

Reziduály zobrazíme funkciou *segments* ako čiary spájajúce dvojice bodov skutočných hodnôt (prvé dva argumenty sú ich súradnice *x* a *y*) a vyrovnaných hodnôt z modelu (ďalšie dva argumenty sú ich súradnice *x* a *y*). Na lepšie odlíšenie im zmeníme farbu na svetlo modrú.

```
> segments(prijem,spotreba,prijem,pred,col="light blue") # zobrazí reziduály ako čiary
```

Ak chceme namiesto reziduálov zobrazit' intervalovú prognózu ako hranice okolo regresnej priamky, najskôr ju musíme vypočítat' (varuje nás, že prognózuujeme pre pôvodné údaje):

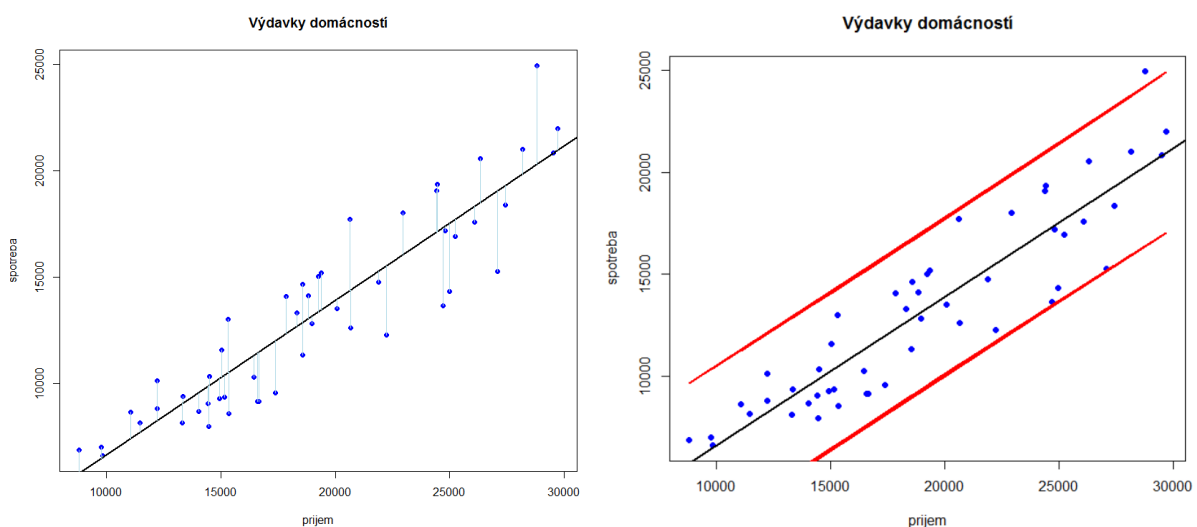
```
> predf <- predict(lin_mod1, interval="prediction", level=0.95) # výpočet intervalu prognózy
```

```
> plot(prijem, spotreba, pch=16, col="blue", main="Výdavky domácností") # graf závislosti
```

```
> abline(lin_mod1, lwd=2) # pridá do grafu regresnú priamku
```

```
> lines(prijem, predf[,2], col="red", lwd=2) # pridá dolnú hranicu ako 2. stĺpec prognózy
```

```
> lines(prijem, predf[,3], col="red", lwd=2) # pridá hornú hranicu ako 3. stĺpec prognózy
```

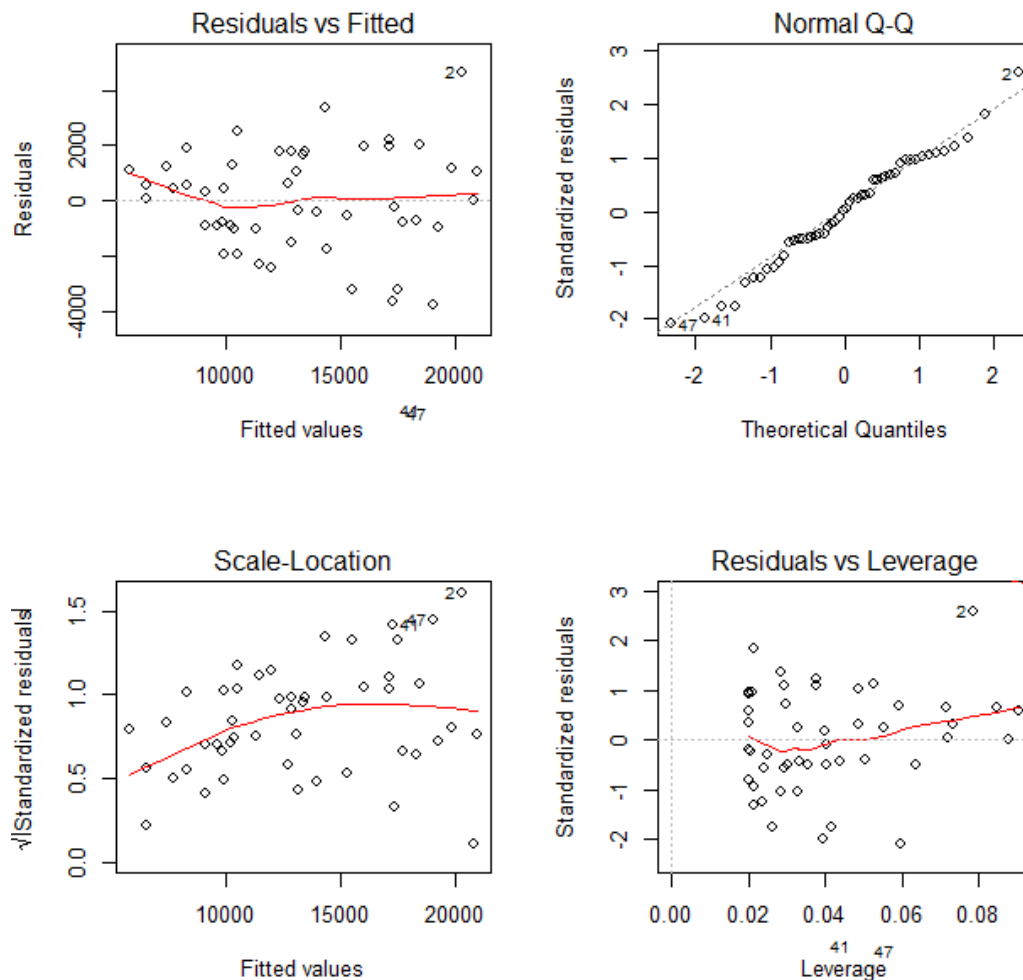


Obrázok 2: Graf regresnej priamky so zobrazenými reziduálmi vľavo a s prognózou vpravo

Medzi generickými funkciami odhadnutého modelu je aj `plot()`, ktorý slúži na zobrazenie diagnostických grafov. Ak by sme priamo zadali `plot(lin_mod1)`, tak by nás program postupne 4 krát vyzval na stlačenie klávesu Enter, aby nám zakaždým zobrazil iný diagnostický graf. Ak chceme zobraziť všetky 4 grafy spolu, musíme pred príkazom `plot` rozdeliť zobrazovaciu plochu na 4 časti, pre každý graf jednu (po zobrazení to vrátíme späť `par(mfrow = c(1, 1))`):

```
> par(mfrow = c(2, 2))
```

```
> plot(lin_mod1)
```



Obrázok 3: Diagnostické grafy odhadnutého lineárneho modelu (4 predvolené typy)

Reziduály podľa prvého grafu vyzerajú byť heteroskedastické (pre väčšie vyrovnané hodnoty sú viac rozptýlené) a podľa druhého QQ grafu majú normálne rozdelenie (body ležia takmer na priamke). Všetky grafy nám do pozornosti dávajú pozorovania 2, 41 a 47, ktoré najviac vplývajú na výsledný odhad.

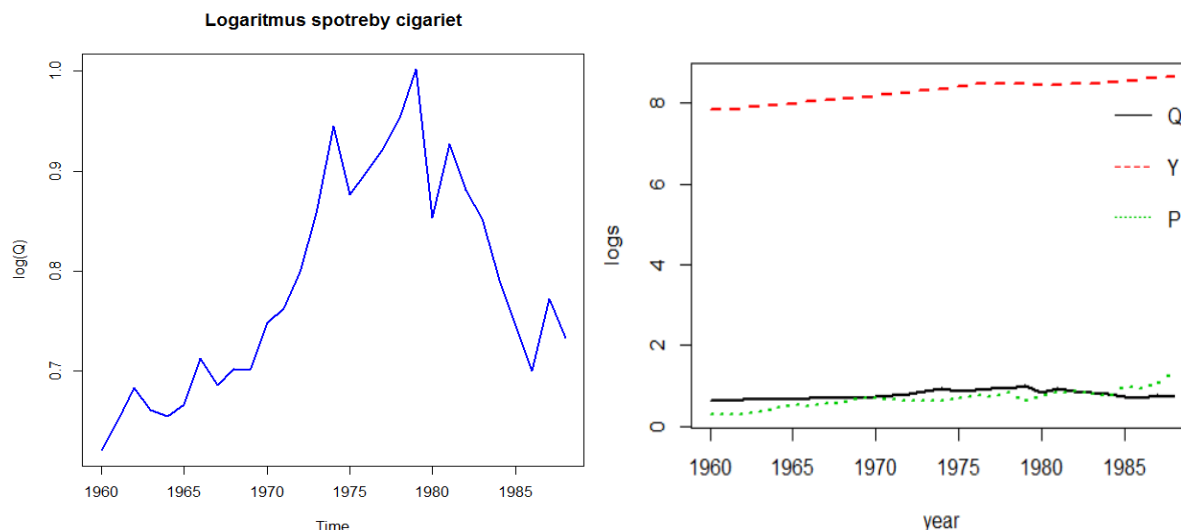
Ak pracujeme s časovými radmi a vytvorili sme údaje typu *ts* (s časovou pečiatkou *time stamp*), môžeme pre každú premennú túto pečiatku zobraziť, napr. pre *Q* príkazom `tsp(Q)`. Zobrazenie časového vývoja jednej premennej, napr. *Q*, sa vykoná príkazom:

```
> plot(Q, col="blue", lwd=2, main="Spotreba cigaret")
```

Zobrazenie časového vývoja viacerých premenných sa vykoná príkazom:

```
> ts.plot(log(Q),log(Y),log(P),gpars=list(xlab="year",ylab="logs",lty=c(1:3),col=1:3))
```

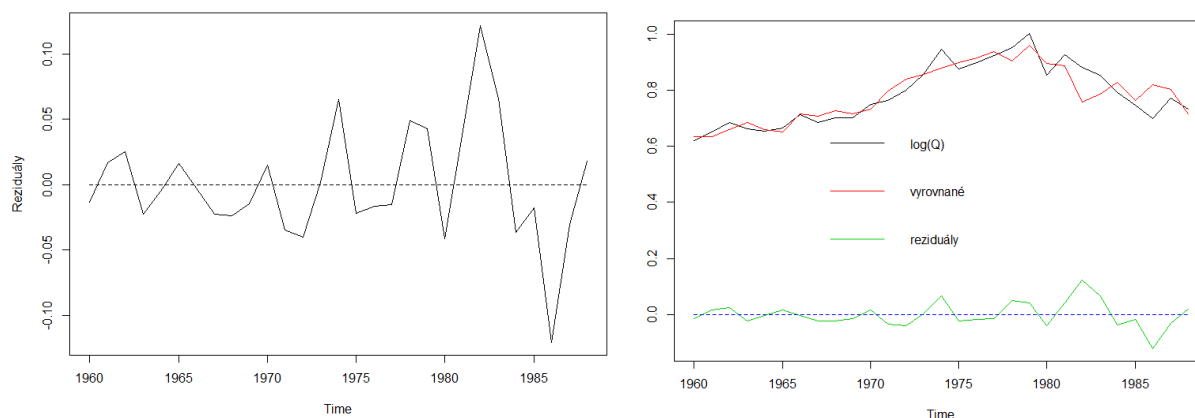
```
> legend("topright", c("Q", "Y", "P"), bty = "n", col = 1:3, lty = c(1:3))
```



Obrázok 4: Graf jednej a viacerých premenných v čase

Rovnako vieme zobrazit' aj priebeh reziduálov alebo porovnať vyrovnané a skutočné hodnoty závislej premennej v čase, ak ich zadáme ako typ *ts*:

```
> res2 <- ts(data=residuals(lin_mod2), start=1960, end=1988, frequency = 1)
> pred2 <- ts(data=predict(lin_mod2), start=1960, end=1988, frequency = 1)
> ts.plot(res2, 0, gpars=list(ylab="Reziduály", lty=c(1,2)))
> ts.plot(log(Q), pred2, res2, 0, gpars=list(lty=c(1,1,1,2), col=1:4))
> legend("center", c("log(Q)", "vyrovnané", "reziduály"), bty = "n", col = 1:3, lty = 1)
```



Obrázok 5: Graf reziduálov v čase a vyrovnané a skutočné hodnoty závislej premennej v čase

5 TESTOVANIE A RIEŠENIE HETEROSKEDASTICITY

Ako sme naznačili už pri diagnostických grafoch, predpokladáme, že reziduály modelu výdavkov domácností na spotrebu nie sú konštantné. Tento predpoklad je dôležité overiť pomocou niektorého z testov heteroskedasticity. Najpoužívanejším testom je Breuschov-Paganov test, ktorý vysvetľuje napr. Lukáčik et al (2013).

Vo všeobecnom lineárnom modeli:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i \quad (4)$$

predpokladáme, že rozptyl náhodných zložiek jednotlivých pozorovaní je funkciou nenáhodných premenných Z_1, Z_2 až Z_m , pričom týmito premennými môžu byť niektoré alebo

aj všetky vysvetľujúce premenné všeobecného modelu X_j , ale aj ľubovoľné iné nestochastické premenné. Zvyčajne predpokladáme lineárny funkčný vzťah medzi rozptylom náhodných zložiek a týmito premennými:

$$\sigma_i^2 = \alpha_0 + \alpha_1 z_{i1} + \alpha_2 z_{i2} + \dots + \alpha_m z_{im} \quad (5)$$

Test overuje nulovú hypotézu o neprítomnosti heteroskedasticity, či sa všetky parametre α_i s výnimkou konštanty súčasne rovnajú 0. Testovacia procedúra sa realizuje v týchto krokoch:

1. odhadneme parametre modelu metódou najmenších štvorcov a vypočítame hodnoty reziduálov e_i a ich štvorce e_i^2 pre každé pozorovanie,
2. vypočítame priemernú hodnotu štvorca reziduálu $\bar{e}^2 = \frac{1}{n} \sum_{i=1}^n e_i^2$,
3. vytvoríme novú pomocnú premennú, ktorej hodnoty získame ako podiel štvorca reziduálu a priemernej hodnoty štvorca reziduálu $g_i = e_i^2 / \bar{e}^2$,
4. odhadneme parametre modelu, kde nová pomocná premenná je závislou premennou a vysvetľujúcimi premennými sú Z_1, Z_2 až Z_m zo vzťahu (5),
5. testovacia štatistika Breuschovho-Paganovho testu je jednou polovicou vysvetleného súčtu štvorcov ESS z tohto modelu.

Za predpokladu, že „problémovou premennou“ spôsobujúcou heteroskedasticitu (Z_1) je príjem, sa B-P postup dá v programovacom systéme R vykonať nasledovnou sekvenciou príkazov:

```
> g <- (residuals(lin_mod1)^2)/(devn=deviance(lin_mod1)/length(residuals(lin_mod1)))
> lin_mod1bp_test <- lm(g ~ prijem) # model závislosti g od „problémových premenných“
> bpstat <- (var(g)*(length(g)-1)-deviance(lin_mod1bp_test))/2 # B-P štatistika
```

Vzhľadom na istú komplikovanosť postupu výpočtu štatistiky B-P navrhli Koenker a Bassett alternatívu tohto testu, ktorá je jednoduchšia a rýchlejšie vypočítateľná. V ich postupe nie je nevyhnutné zavádzať pomocnú premennú do testovacieho modelu, ale postačujú priamo štvorce reziduálov získané metódou najmenších štvorcov. Testovacia štatistika Koenkerovho verzie testu sa vypočíta ako nR^2 .

```
> lin_mod1bp_koenker <- lm(residuals(lin_mod1)^2 ~ prijem) # testovací model Koenkera
> koenker <- summary(lin_mod1bp_koenker)
> bpkoenker <- length(residuals(lin_mod1bp_koenker))*koenker$r.squared # B-P Koenker
```

Nulovú hypotézu homoskedasticity pre obe štatistiky zamietame, ak sú väčšie ako kritická hodnota χ^2 -rozdelenia s počtom stupňov voľnosti = počtu premenných testovacieho modelu. Štatistiky $bpstat = 6,0849$ aj $bpkoenker = 7,2023$ sú väčšie ako $\chi^2(1) = 3,8415$.

Vidíme, že v celom postupe sa odvolávame iba na vypočítané hodnoty z pôvodného modelu. A takisto, že zmeniť „problémovú premennú“ spôsobujúcu heteroskedasticitu (resp. pridať ich viac) je veľmi jednoduché, lebo len v druhom príkaze inak špecifikujeme pomocný model závislosti pomocnej premennej g . Vytvárať novú funkciu, aj keď je to lákavé, nemusíme, lebo už bola vytvorená, ale aby sme ju mohli používať, vyžaduje inštalovať prídavný balík *lmtest*.

Po nainštalovaní a sprístupnení balíka *lmtest* príkazmi:

```
> install.packages("lmtest")
> library(lmtest)
```

sa Koenkerova verzia Breuschovho-Paganovho testu heteroskedasticity spustí pre odhadnutý model príkazom:

```
> bptest(lin_mod1)
```

Pokiaľ sa zadá ako argument iba názov modelu, sú za „problémové premenné“ považované všetky premenné modelu, ktorých reziduály podozrievame z heteroskedasticity, čo je náš prípad (takou premennou je *prijem*). Ak chceme testovať vlastnú predpokladanú závislosť reziduálov, zadáme ju v tvare odhadovaného modelu ako druhý argument. Potom:

```
> bptest(lin_mod1, ~ prijem + I(prijem^2))
```

testuje závislosť v tvare: $e_i^2 = \alpha_0 + \alpha_1 \text{prijem}_i + \alpha_2 \text{prijem}_i^2$, pričom sme využili možnosť izolácie výpočtu kvadratickej funkcie pomocou $I()$ ako *insulate* v zadaní lineárneho modelu. Táto špecifikácia, mimochodom, zodpovedá Whiteovmu testu závislosti štvorcov reziduálov od vysvetľujúcich premenných, štvorcov vysvetľujúcich premenných a vzájomných súčinov vysvetľujúcich premenných (v modeli s 1 vysvetľujúcou premennou vzájomné súčiny nie sú).

Riešením heteroskedasticity je odhad váženou metódou najmenších štvorcov, čo znamená, že obe strany všeobecného lineárneho modelu (4) vydělíme štandardnou odchýlkou a dostaneme:

$$\frac{y_i}{\sigma_i} = \beta_0 \frac{1}{\sigma_i} + \beta_1 \frac{x_{i1}}{\sigma_i} + \beta_2 \frac{x_{i2}}{\sigma_i} + \dots + \beta_k \frac{x_{ik}}{\sigma_i} + \frac{u_i}{\sigma_i} \quad (6)$$

čo už je model s homoskedastickými náhodnými zložkami. Pre náš prípad podľa B-P testu, kde štvorce reziduálov (nahrádzajúce rozptyl) sú funkciou premennej *prijem*, to znamená, že celú rovnicu predelíme odmocninou premennej *prijem*:

$$\frac{\text{spotreba}_i}{\sqrt{\text{prijem}_i}} = \beta_0 \frac{1}{\sqrt{\text{prijem}_i}} + \beta_1 \frac{\text{prijem}_i}{\sqrt{\text{prijem}_i}} + \frac{u_i}{\sqrt{\text{prijem}_i}} \quad (7)$$

Pri odhade modelu využívame nielen izoláciu výpočtu druhej odmocniny, ale aj odhad bez absolútneho člena, ktorý sa v špecifikácii zadá pridaním -1 na koniec (alebo 0 + na začiatok):

```
> lin_mod1_vaz <- lm(I(spotreba/sqrt(prijem)) ~ I(1/sqrt(prijem)) + I(prijem/sqrt(prijem)) - 1)
> summary(lin_mod1_vaz) # komplexnejší výpis odhadu váženého modelu
```

Funkcia *lm* má možnosť pri odhade lineárneho modelu priamo zadať váhy prostredníctvom argumentu *weights*. Z premennej zadanej pri odhade váženou metódou ako váhy sa automaticky počíta druhá odmocnina a až tá sa použije pri výpočte, preto sa zadá ako $1/\text{prijem}$:

```
> lin_mod1_weight <- lm(spotreba ~ prijem, weights = I(1/(prijem)))
> summary(lin_mod1_weight)
```

Výsledky odhadu parametrov oboch spôsobov sú rovnaké, sprievodné štatistiky sa líšia.

6 TESTOVANIE A RIEŠENIE AUTOKORELÁCIE

Predpokladajme, že reziduály modelu spotreby cigariet sú autokorelované. Tento predpoklad je dôležité overiť pomocou niektorého z testov autokorelácie. Najpoužívanejším testom autokorelácie 1. rádu je Durbinov-Watsonov test, ktorý vysvetľuje napr. Lukáčik et al (2013).

Vytvárať funkciu na výpočet Durbinovej-Watsonovej štatistiky nemusíme, lebo už bola vytvorená, a rovnako ako v predošlom prípade vyžaduje inštalovať prídavný balík *lmtest*. Keďže to sme už urobili, stačí nám zadať príkaz *dwttest()* s argumentom názvom modelu:

```
> dwttest(lin_mod2) # DW test autokorelácie rádu 1
```

Výsledná hodnota $DW = 1,4569$ je doplnená pravdepodobnostnou hodnotou $p\text{-value} = 0,0199$, ktorá zobrazuje hladinu významnosti, od ktorej zamietame H_0 o neprítomnosti autokorelácie.

Pri publikovanej pravdepodobnostnej hodnote nemusíme konfrontovať výsledok s tabuľkami, ktoré obsahujú oblasť na ktorej nevieme rozhodnúť (sem patrí výsledok), ale preferujeme záver získaný z *p-value*, lebo tá zohľadňuje konkrétnu maticu vysvetľujúcich premenných.

Na testovanie autokorelácie, a to nielen 1. rádu, existujú ďalšie testovacie postupy navrhnuté Ljungom a Boxom alebo Breuschom a Godfreyom. Ljungov-Boxov test vypočítavaný pre ľubovoľný časový rad zadáme príkazom *Box.test()* s argumentmi rad reziduálov a typ:

```
> Box.test(residuals(lin_mod2), type = "Ljung-Box") # LB test autokorelácie rádu 1 (default)
> Box.test(residuals(lin_mod2), type = "Ljung-Box", lag = 2) # LB test autokorelácie rádu 2
```

Výsledná hodnota $LB_1 = 2,2892$ je menšia ako kritická hodnota χ^2 -rozdelenia s počtom stupňov voľnosti = testovanému rádu autokorelácie, $\chi^2(1) = 3,8415$. Nezamietame H_0 o neprítomnosti autokorelácie 1. rádu (ani 2. rádu, lebo $LB_2 = 3,8163 < \chi^2(2) = 5,9915$).

Breuschov a Godfreyho test pre reziduály zadáme príkazom *bgtest()* s názvom modelu:

```
> bgtest(lin_mod2) # BG test autokorelácie rádu 1 (default)
> bgtest(lin_mod2, order = 2) # BG test autokorelácie rádu 2
```

Výsledná hodnota $BG_1 = 2,8854$ je menšia ako kritická hodnota χ^2 -rozdelenia s počtom stupňov voľnosti = testovanému rádu autokorelácie, $\chi^2(1) = 3,8415$. Nezamietame H_0 o neprítomnosti autokorelácie 1. rádu (ani 2. rádu, lebo $BG_2 = 4,8185 < \chi^2(2) = 5,9915$).

Ak by sme postup zopakovali pre model spotreby cigariet bez umelej premennej D82:

$$\ln Q_t = \beta_0 + \beta_1 \ln Y_t + \beta_2 \ln P_t + u_t \quad (8)$$

tak všetky testy zamietnu neprítomnosť autokorelácie 1. rádu:

```
> lin_mod2b <- lm(log(Q) ~ log(Y) + log(P)) # odhad zapíšeme do objektu lin_mod2b
> summary(lin_mod2b) # komplexnejší výpis informácií o parametroch z odhadu
> dwtest(lin_mod2b) # DW test autokorelácie rádu 1
> Box.test(residuals(lin_mod2), type = "Ljung-Box") # LB test autokorelácie rádu 1
> bgtest(lin_mod2) # BG test autokorelácie rádu 1
```

lebo $DW = 1,0201$ s jej *p-value* = 0,0006 a $LB_1 = 7,2384$ a $BG_1 = 8,9685$ sú väčšie ako $\chi^2(1)$.

Ako riešenie autokorelácie zvolíme dynamizáciu modelu (aj keď podľa dvoch testov pomohlo zaradenie umelej premennej – vidíme v modeli (2)), tzn. pridanie oneskorenej závislej premennej medzi vysvetľujúce premenné (keďže máme autokoreláciu rádu 1, pridáme závislú premennú oneskorenú o 1 obdobie) a logaritmický model spotreby cigariet bude mať tvar:

$$\ln Q_t = \beta_0 + \beta_1 \ln Y_t + \beta_2 \ln P_t + \beta_3 \ln Q_{t-1} + u_t \quad (9)$$

Odhad parametrov tohto modelu si vyžaduje vytvorenie nového dátového rámca, do ktorého zahrnieme časový prienik existujúcich použitých premenných (oneskorená premenná existuje od 1961 do 1989, pričom všetky ostatné premenné od 1960 do 1988) od 1961 do 1968. Príkaz na prienik údajov je *ts.intersect*. Pri vytváraní prieniku využijeme funkciu *lag* na oneskorenie objektu *ts* o 1 obdobie, $k = -1$. Nový dátový rámec použijeme na odhad modelu *lin_mod3*.

```
> ram_dat <- ts.intersect(lQ = log(Q), lQ1 = lag(lQ, k = -1), lY = log(Y), lP = log(P))
> lin_mod3 <- lm(lQ ~ lY + lP + lQ1, data = ram_dat)
> summary(lin_mod3)
```

Testovaním autokorelácie reziduálov modelu *lin_mod3* sa nepreukáže jej prítomnosť.

7 PRÍDAVNÉ BALÍKY PRE ĎALŠIE EKONOMETRICKÉ ANALÝZY

Tento úvod nemôže postihnúť rozsiahle možnosti, ktoré ponúka ekonometrická analýza. Ale prakticky pre každý bežný problém už boli navrhnuté prídavné balíky.

Problematika dynamických modelov s oneskorenými premennými sa lepšie rieši pomocou balíka *dynlm*, v ktorom sú dovolené diferencie premennej pomocou *d(meno_premennej)* alebo oneskorenia premennej o *k* období pomocou *L(meno_premennej, k)*.

Oblasť panelových dát je spracovaná v balíku *plm*, kde sa v odhade modelu navyše špecifikuje typ modelu *model = "pooling"*, *model = "within"* pre fixné efekty alebo *model = "random"* pre náhodné efekty so spresnením cez *random.method*. Samozrejme v dátovom rámci je charakterizovaný prierezný aj časový rozmer, napr. pomocou *index = c("country", "year")*.

Pre viacrovnicový systém lineárnych rovníc je určený balík *systemfit*, ktorý umožňuje odhady metódami "OLS", "SUR", "2SLS", "3SLS", "WLS" a "W2SLS".

Pre regresné modely priestorovej ekonometrie sú určené balíky *spdep* a *sphet* (prístup GMM). Balík *splm* je vhodný pre priestorové panely a probit modely sú v balíku *spatialprobit*.

Počet publikovaných balíkov neustále rastie a mnoho oblastí je riešených viacerými balíkmi, preto na uľahčenie orientácie medzi nimi vytvorili tvorcovia systematizovaný katalóg balíkov podľa oblastí na stránke <https://cran.r-project.org/web/views/>, kde je zaradená aj ekonometria <https://cran.r-project.org/web/views/Econometrics.html>.

Použitá literatúra

1. ALBERT, J., RIZZO, M. 2012. *R by Example*. New York: Springer, 2012.
2. CRAWLEY, M. J. 2013. *The R Book, 2nd. ed.* Chichester: John Wiley & Sons, 2013.
3. FARAWAY, J. J. 2015. *Linear Models with R, 2nd ed.* Boca Raton: Taylor & Francis, 2015.
4. HATRÁK, M. 2007. *Ekonometria*. Bratislava: IURA Edition, 2007.
5. KLEIBER, Ch., ZEILEIS, A. 2008. *Applied Econometrics with R*. New York: Springer, 2008.
6. LUKÁČIK, M., LUKÁČIKOVÁ, A., SZOMOLÁNYI, K. 2013. *Ekonometria 1*. Bratislava: Vydavateľstvo EKONÓM, 2013.
7. LUKÁČIK, M., LUKÁČIKOVÁ, A., SZOMOLÁNYI, K. 2016. *Úvod do práce so systémom R*. In: Zborník mezinárodného vedeckého seminára Nové trendy v ekonometrii a operačným výzkumu, Bratislava: Vydavateľstvo EKONÓM, 2016.
8. RAMANATHAN, R. 1998. *Introductory Econometrics with Applications, 4th ed.* Harcourt Brace College, 1998.
9. RODRIGUES, B. 2014. *Introduction to Programming Econometrics with R*. Dostupné na: <http://qe4policy.ec.unipi.it/wp-content/uploads/2015/10/Introduction-to-programming-Econometrics-with-R-Draft.pdf> [20.10.2016]
10. VENABLES, W. N., SMITH, D. M., the R Core Team. 2016. *An Introduction to R Notes on R: A Programming Environment for Data Analysis and Graphics*. Dostupné na: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf> [31.10.2016]

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

ÚVOD DO PRÁCE SO SYSTÉMOM R¹

INTRODUCTION TO R SYSTEM

Martin Lukáčik, Adriana Lukáčiková, Karol Szomolányi

Abstrakt

Rozvoj štatistickej 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 základy práce s programovacím systémom R, ktorý je voľne dostupný a predstavuje kvalitnú alternatívu komerčným špecializovaným softvérovým produktom.


Kľúčové slová: analýza, programovací systém R, RStudio

Abstract

The development of statistical 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 the basics of R programming system, which is freely available and it is a quality alternative to commercial software product.

Keywords: analysis, R programming system, RStudio

1 INŠTALÁCIA SYSTÉMU R A JEHO NADSTAVBA RSTUDIO

Inštalácia programovacieho systému R v prostredí Windows je jednoduchá. Z internetovej stránky <https://cran.r-project.org/bin/windows/base/> stiahneme inštalačný program R. Spustíme ho a zvolíme jazyk, ktorý sa bude používať počas inštalácie (napr. čeština). V jednotlivých krokoch v sprievodcovi inštalácie potvrdíme, že rozumieme licencií a zvolíme priečinok, do ktorého bude systém R nainštalovaný. Posledným dôležitým krokom je výber súčastí, ktoré sa majú nainštalovať, kde si môžeme vybrať medzi 32 a 64 bitovou verziou programu. Odporúča sa inštalovať 32 bitovú verziu programu v 32 bitovom operačnom systéme a 64 bitovú verziu programu v 64 bitovom operačnom systéme Windows. Akú verziu operačného systému máme nainštalovanú, zistíme v okne, ktoré spustíme súčasným stlačením klávesov  a Pause.

Po inštalácii programovacieho systému R odporúčame nainštalovať aj príjemnejšie grafické prostredie k tomuto systému – RStudio, ktoré nájdeme na internetovej stránke <https://www.rstudio.com/products/rstudio/download/>. Inštalácia si opäť vyžaduje iba voľbu priečinku, do ktorého bude RStudio nainštalované. Všetky príkazy už zadávame iba v jeho prostredí cez konzolu.

RStudio (rovnako ako aj R bez neho) využíva dva základné spôsoby zadávania príkazov, a to buď pomocou konzoly, v ktorej sa po zadaní Enter na konci príkazu, tento ihneď vykoná alebo pomocou tzv. skriptu teda sekvencie príkazov, ktorá sa ukladá do súborov s príponou .R (iba veľké R). Základná obsluha programu je umožnená aj pomocou hlavnej ponuky.

¹ Príspevok vznikol s podporou projektu KEGA 044EU-4/2015 "Vypracovanie monografie Bayesovská ekonometria pre prax a študentov postgraduálneho štúdia".

2 JEDNODUCHÉ PRIRADENIE HODNÔT

Na začiatku práce po spustení programu alebo automaticky pri spustení nového projektu (*File>New_Project*) sa nastavuje pracovný priečinok, aby program vedel odkiaľ má získavať údaje. Nastavíme ho pomocou ponuky *Session>Set_Working_Directory>Choose_Directory*. Ak si vyberieme napr. priečinok *FHI_EU/Documents* na disku Z, program túto zmenu vykoná a v konzole zároveň vidíme zobrazený príkaz a jeho štruktúru (syntax):

```
> setwd("Z:/FHI_EU/Documents")
```

Pri práci so systémom R je dôležité poznať štruktúru jednotlivých príkazov. Pomoc pri štruktúre príkazu s vysvetlením a ukázkovými príkladmi získame cez zadanie otáznika a názvu daného príkazu, napríklad:

```
> ?mean
```

ktorý zobrazí pomoc k príkazu *mean* slúžiacemu na výpočet aritmetického priemeru v okne prehliadača. Analogicky môžeme pomoc získať cez zadanie názvu daného príkazu v zátvorkách za príkazom *help*, napríklad:

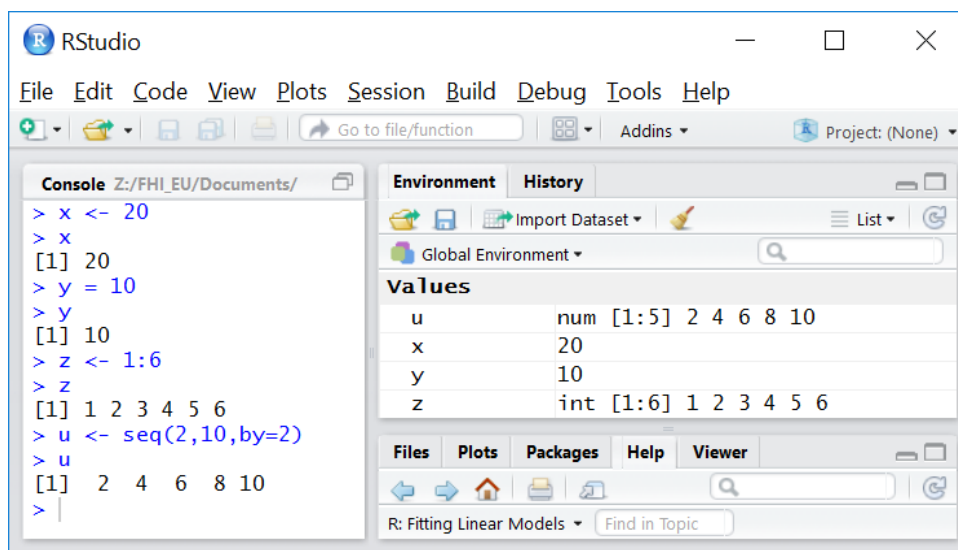
```
> help(lm)
```

ktorý zobrazí pomoc k príkazu *lm* slúžiacemu na odhad parametrov lineárneho modelu v okne prehliadača.

Aj z príkladov, ktoré sú zobrazené pri vysvetleniach týchto príkazov, vidíme, akým spôsobom sa pracuje s hodnotami v programe. Operátorom priradenia je `<-` (znak menší a pomlčka vytvárajúce šípku priradenia) alebo kvôli konvencii aj `=` (znak rovnosti).

Aritmetická postupnosť hodnôt s krokom 1 sa zadá pomocou operátora `:` (dvojbodka), s iným krokom ako 1 pomocou funkcie *seq*, za ktorú sa do zátvoriek zadajú argumenty prvá hodnota, posledná hodnota a krok zadaný ako *by=číslo* oddelené čiarkami. Zadaním názvu objektu sa zobrazí jeho zadaný obsah. Text písaný za znak mriežky `#` (*hashtag*) predstavuje našu textovú poznámku k príkazu.

```
> x <- 20 # priradenie číselnej hodnoty 20 do x
> y = 10 # priradenie číselnej hodnoty 10 do y
> z <- 1:6 # priradenie číselnej postupnosti od 1 do 6 do z
> u <- seq(2,10,by=2) # priradenie postupnosti po dvoch od 2 do 10 do u
```

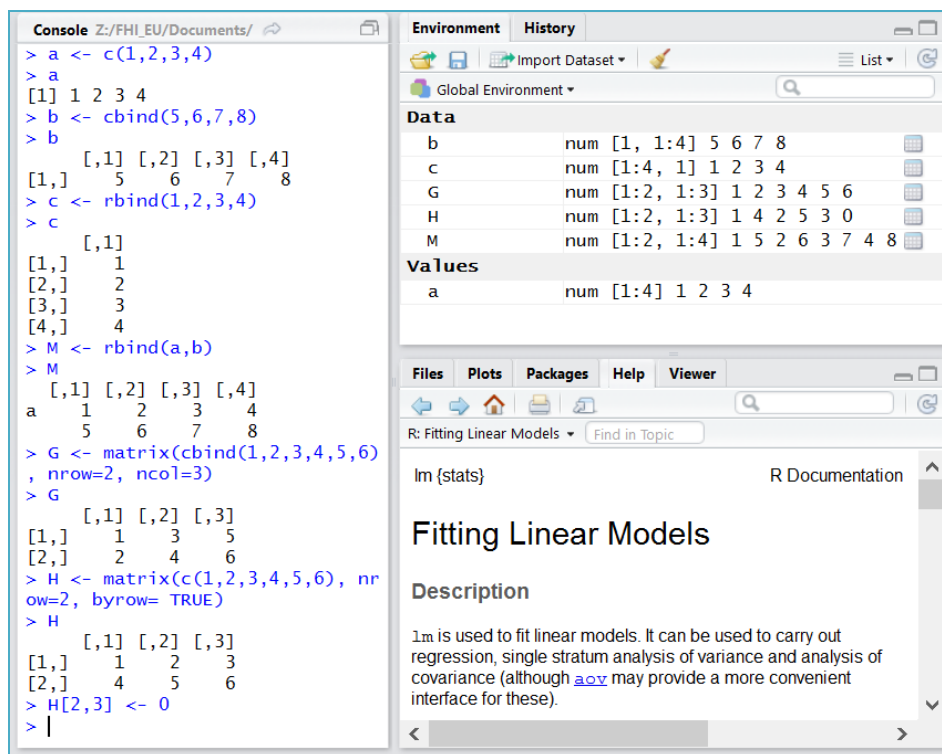


Obrázok 1: Jednoduché priradenia hodnôt v systéme R

3 VEKTORY, MATICE A ČÍSELNÉ ZOZNAMY

Zoznam hodnôt pripomínajúci vektor sa vytvorí pomocou funkcie `c`, za ktorú sa do zátvoriek zadá zoznam jeho prvkov oddelených čiarkami. Riadkový vektor v matematickom zmysle sa vytvorí pomocou funkcie `cbind`, za ktorú sa do zátvoriek zadá zoznam jeho prvkov oddelených čiarkami. Maticu z riadkových vektorov (zoznamov) vytvoríme pomocou funkcie `rbind` (slúžiacej na vytvorenie stĺpcového vektora), za ktorú do zátvoriek zadáme zoznam vektorov oddelených čiarkami. Rozmery matice alebo vektora zobrazíme pomocou funkcie `dim`, za ktorú zadáme do zátvoriek názov objektu. Matica sa všeobecne vytvorí funkciou `matrix`, ktorej argumentom je vektor alebo zoznam hodnôt a jej rozmery sa špecifikujú pomocou argumentov `nrow` a/alebo `ncol`. Údaje sa v matici vyplňajú po stĺpcoch. Ak ich chceme zadávať po riadkoch, musíme nastaviť argument `byrow` funkcie `matrix` na hodnotu `TRUE`. Zobrazenie alebo prepis konkrétneho prvku matice získame zápisom jeho pozície v hranatých zátvorkách za jej názvom.

```
> a <- c(1,2,3,4) # vytvorenie zoznamu hodnôt a
> b <- cbind(5,6,7,8) # vytvorenie riadkového vektora b
> c <- rbind(1,2,3,4) # vytvorenie riadkového vektora c
> M <- rbind(a,b) # vytvorenie matice M z riadkových zoznamov a vektorov
> G <- matrix(cbind(1,2,3,4,5,6), nrow=2, ncol=3) # vytvorenie matice G
> H <- matrix(c(1,2,3,4,5,6), nrow=2, byrow=TRUE) # vytvorenie matice H
> H[2,3] <- 0 # prepis prvku v druhom riadku a treťom stĺpci matice H na 0
```



Obrázok 2: Vytváranie vektorov a matíc v systéme R

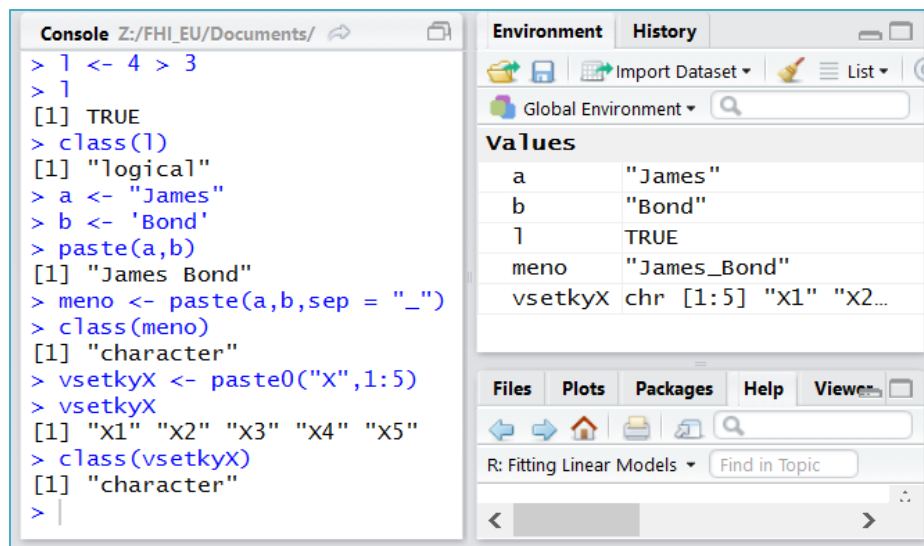
Z obrázkov 1 a 2 vidíme, že v pracovnom priestore (*Environment*) sa vytvárajú jednotlivé objekty pre hodnoty (*Values*) alebo dáta (*Data*) podľa daných príkazov. Ak chceme vymazať pracovný priestor, zvolíme voľbu `Session > Clear Workspace`. Všetky hodnoty z neho sa dajú v ponuke `Session` uložiť do súboru s príponou `.RData` alebo z neho späť nahráť, čo RStudio ponúka aj pri ukončení práce. Ak chceme vyčistiť zápisy v konzole, zvolíme `Edit > Clear Console`. Napriek ich vyčisteniu sa klávesmi šípka nahor a nadol dokážeme pohybovať medzi jednotlivými zadanými príkazmi, ktoré sa zobrazujú po riadkoch.

Takisto po ukončení práce (*Quit Session*) zostáva v pracovnom priečinku uložený súbor *.Rhistory*, v ktorom sú uložené všetky príkazy (aj nesprávne zadané) postupne použité počas práce v konzole.

4 REŤAZCE A ICH SPÁJANIE

Okrem objektov, ktoré sú celočíselného typu (*integer*), alebo majú plávajúcu desatinnú čiarku (*numeric*), prípadne sú maticového typu (*matrix*), systém rozlišuje aj objekty logického typu (*logical*) nadobúdajúce hodnoty *TRUE* (skrátene *T*) alebo *FALSE* (skrátene *F*) a takisto reťazce (*character*). Typ objektu zistíme príkazom *class* s argumentom názov objektu. Reťazec zadáme do objektu ako ľubovoľnú postupnosť znakov v úvodzovkách alebo apostrofoch. Príkaz *paste* spája reťazce do jedného celku, pričom oddeľovač je určený argumentom *sep*. Ak zvolíme príkaz *paste0*, spoja sa bez oddeľovača.

```
> b <- "James" # priradí do objektu b reťazec James
> c <- 'Bond' # priradí do objektu c reťazec Bond
> meno <- paste(b,c,sep = "_") # spojí b a c znakom _ do objektu meno
> class(meno) # zistí typ objektu meno
> vsetkyX <- paste0("X",1:5) # vytvorí zoznam X1 až X5
```



Obrázok 3: Reťazce a ich spájanie v systéme R

5 ARITMETICKÉ A LOGICKÉ OPERÁTORY

Základnými aritmetickými operátormi sú:

+ pre sčítanie, - pre odčítanie, * pre násobenie, / pre delenie, ^ pre umocňovanie, %/% pre celočíselné delenie (div) a %% pre zvyšok po celočíselnom delení (mod).

Logickými operátormi sú:

== pre rovná sa, != pre nerovná sa, <, >, <=, >= pre porovnania, && pre logické a (and), // pre logické alebo (or), ! pre negáciu a *is.na(x)* pre zistenie, či má x chýbajúce hodnoty.

Ak matice **X** a **Y** majú potrebné rozmery pre zodpovedajúce operácie:

$X \%*\% Y$ vypočíta súčin matic, $X * Y$ vypočíta súčin matic po prvkoch, $X \%x\% Y$ (takisto aj *kroncker(X, Y)*) vypočíta Kroneckerov súčin matic, *t(X)* zobrazí transponovanú maticu

k matici X , $diag(X)$ zobrazí diagonálne prvky matice X , $eigen(X)$ vypočíta vlastné čísla a charakteristické vektory matice X a $chol(X)$ vypočíta Choleskeho dekompozíciu matice X .

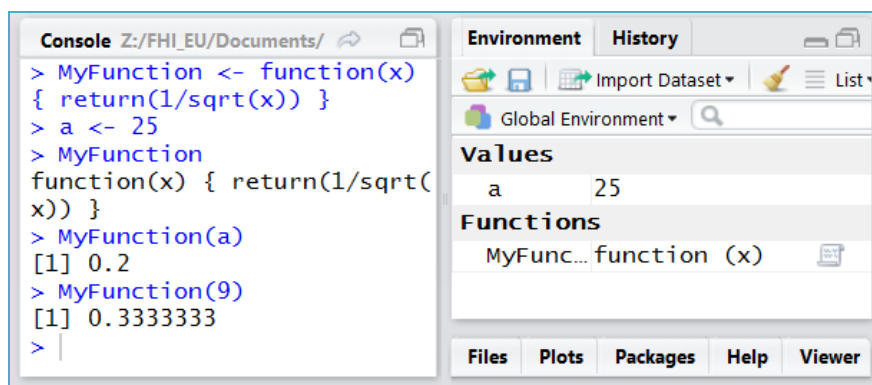
6 MATEMATICKÉ, ŠTATISTICKÉ A VLASTNÉ FUNKCIE

Aj keď programovací systém R slúži najmä na vytváranie vlastných funkcií, napr. Rodrigues (2014) ukazuje jednoduchú deklaráciu $f(x) = 1/\sqrt{x}$:

```
> MyFunction <- function(x) { return(1/sqrt(x)) } # vráti inverziu odmocniny
```

tak obsahuje už v základnej inštalácii základné matematické a štatistické funkcie:

$abs(x)$, $sqrt(x)$, $round(x,digits=n)$, $exp(x)$, $log(x)$, $log10(x)$, $sin(x)$, $cos(x)$, $tan(x)$, $factorial(x)$, $sum(x)$, $min(x)$, $max(x)$, $mean(x)$, $var(x)$, $sd(x)$, $cov(x, y)$, $cor(x, y)$, $median(x)$, $quantile(x, p)$, $table(x)$ a $summary(x)$, ktorých funkcionalita je zrejmá z názvu funkcie.



Obrázok 4: Vytvorenie vlastnej funkcie v systéme R

Systém R ponúka viaceré funkcie pre pravdepodobnostné rozdelenia: $dnorm(x)$ pre hodnotu funkcie hustoty normovaného normálneho rozdelenia, $pnorm(q)$ pre hodnotu distribučnej funkcie normovaného normálneho rozdelenia pre kvantil q , $qnorm(p)$ vráti hodnotu kvantilu pre percentil p a $rnorm(n)$ vráti n náhodných čísel z normovaného normálneho rozdelenia.

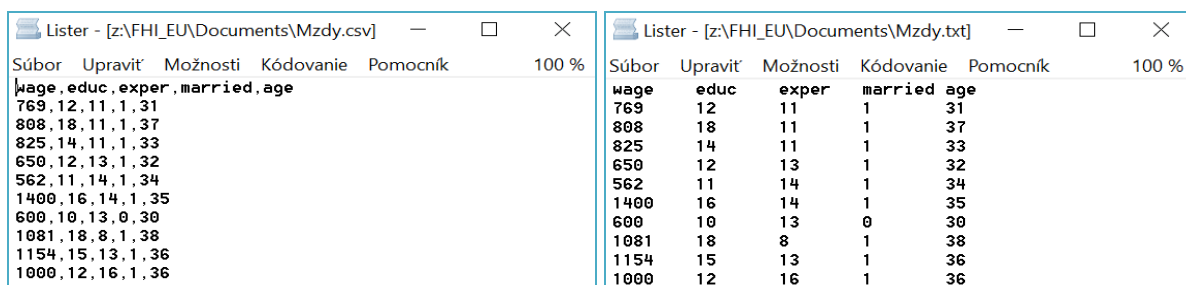
Obdobné funkcie začínajúce písmenami d , p , q a r pre funkciu hustoty, distribučnú funkciu, kvantil a náhodné čísla existujú aj pre ďalšie spojité rozdelenia, pričom vypíšeme len funkcie slúžiace na získanie hodnôt hustoty: $dnorm(x, mean, sd)$ pre normálne rozdelenie, $dt(x, df)$ pre Studentovo rozdelenie, $dcauchy(x, location, scale)$ pre Cauchyho rozdelenie, $dlogis(x, location, scale)$ pre logistické rozdelenie, $dexp(x, rate)$ pre exponenciálne rozdelenie, $df(x, df1, df2)$ pre Fisherovo rozdelenie, $dchisq(x, df)$ pre χ^2 rozdelenie, $dlnorm(x, mean, sd)$ pre lognormálne rozdelenie, $dweibull(x, shape, scale)$ pre Weibullovo rozdelenie, $dgamma(x, shape, rate)$ pre gama rozdelenie, $dunif(x, min, max)$ pre rovnomerné rozdelenie na intervale (min, max) a $dbeta(x, a, b)$ pre beta rozdelenie na intervale (a, b) .

Pre diskrétna rozdelenia je situácia podobná, existujú funkcie začínajúce písmenami d , p , q a r pre pravdepodobnostnú funkciu, distribučnú funkciu, kvantil a náhodné čísla, pričom vypíšeme len funkcie slúžiace na získanie hodnôt pravdepodobností: $dbinom(x, n, p)$ pre binomické rozdelenie, $dgeom(x, prob)$ pre geometrické rozdelenie, $dnbinom(x, size, prob)$ pre negatívne binomické rozdelenie, $dhyper(x, m, n, k)$ pre hypergeometrické rozdelenie, $dpois(x, lambda)$ pre Poissonovo rozdelenie, $dmultinom(x, mean, sd)$ pre multinomické rozdelenie, $dsignrank(x, n)$ pre rozdelenie Wilcoxonovej jednovýberovej štatistiky a $dwilcox(x, m, n)$ pre rozdelenie Wilcoxonovej dvojjvýberovej štatistiky.

7 IMPORT ÚDAJOV DO SYSTÉMU R

Pred analýzou ľubovoľných dát ich najskôr musíme importovať do systému R. Predpokladajme, že dáta sú uložené v pracovnom priečinku ako *Mzdy.csv*, teda súbor vo formáte hodnôt oddelených čiarkou (*Comma-separated value*), ktorý sa dá exportovať z rôznych programov alebo uložiť v Exceli. Súbor v tomto tvare vidíme na obrázku 5 vľavo.

Pripomíname, že v USA je desatinným oddeľovačom bodka a hodnoty sú oddelené čiarkou, kým na Slovensku je desatinným oddeľovačom čiarka a hodnoty sú oddelené v súbore typu *.csv* bodkočiarkou.



Súbor	Upraviť	Možnosti	Kódovanie	Pomocník	100 %
wage,educ,exper,married,age					
769,12,11,1,31					
808,18,11,1,37					
825,14,11,1,33					
650,12,13,1,32					
562,11,14,1,34					
1400,16,14,1,35					
600,10,13,0,30					
1081,18,8,1,38					
1154,15,13,1,36					
1000,12,16,1,36					

Súbor	Upraviť	Možnosti	Kódovanie	Pomocník	100 %
wage	educ	exper	married	age	
769	12	11	1	31	
808	18	11	1	37	
825	14	11	1	33	
650	12	13	1	32	
562	11	14	1	34	
1400	16	14	1	35	
600	10	13	0	30	
1081	18	8	1	38	
1154	15	13	1	36	
1000	12	16	1	36	

Obrázok 5: Súbor typu *.csv* a súbor s údajmi zarovnanými v stĺpcoch

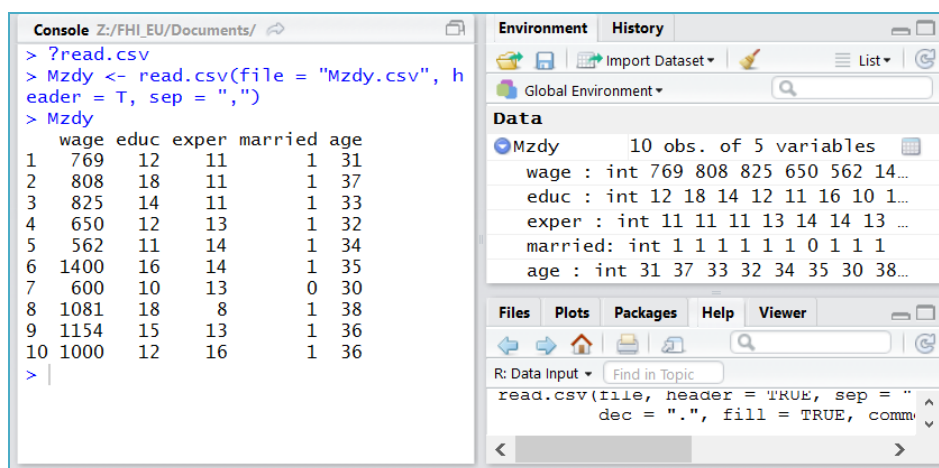
Súbor typu *.csv* importujeme do systému R príkazom *read.csv*, ktorého argumenty *file* špecifikujú názov súboru, *header* či sú v prvom riadku názvy premenných, *sep* aký znak sa používa na oddelenie čísel, *dec* aký znak je desatinný oddeľovač a *skip* koľko riadkov od začiatku treba preskočiť, kým sa začnú načítavať hodnoty (ak je nad údajmi a názvami nejaký popisný text).

```
> Mzdy <- read.csv(file = "Mzdy.csv", header = T, sep = ",") # dáta v US
```

Príkaz na import dát pre údaje uložené podľa slovenských pravidiel vyzerá takto:

```
> Mzdy <- read.csv(file = "Mzdy.csv", header = T, sep = ";", dec = ",")
```

Druhou možnosťou je, že dáta sú uložené v pracovnom priečinku ako *Mzdy.txt*, teda obyčajný textový súbor s údajmi zarovnanými v stĺpcoch pod sebou. Súbor v tomto tvare vidíme na obrázku 5 vpravo.



```
> ?read.csv
> Mzdy <- read.csv(file = "Mzdy.csv", header = T, sep = ",")
> Mzdy
  wage educ exper married age
1  769  12   11      1    31
2  808  18   11      1    37
3  825  14   11      1    33
4  650  12   13      1    32
5  562  11   14      1    34
6 1400  16   14      1    35
7   600  10   13      0    30
8 1081  18    8      1    38
9 1154  15   13      1    36
10 1000  12   16      1    36
> |
```

Environment History

Global Environment

Data

Mzdy 10 obs. of 5 variables

wage : int 769 808 825 650 562 14...

educ : int 12 18 14 12 11 16 10 1...

exper : int 11 11 11 13 14 14 13 ...

married: int 1 1 1 1 1 1 0 1 1 1

age : int 31 37 33 32 34 35 30 38...

R: Data Input

```
read.csv(file, header = TRUE, sep = ",",
dec = ".", fill = TRUE, comm
```

Obrázok 6: Import údajov zo súboru typu *.csv* do systému R

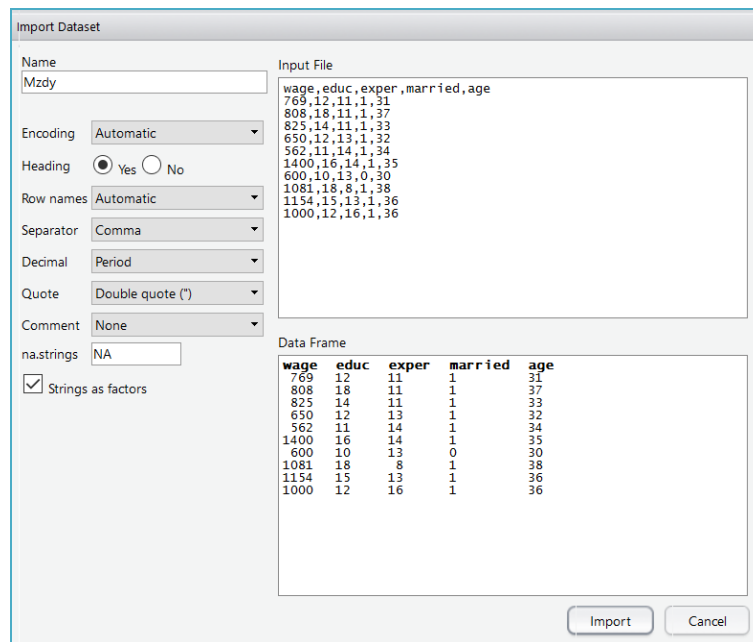
V prípade obyčajnej tabuľky v textovom súbore použijeme príkaz *read.table*, ktorý má rovnaké argumenty ako príkaz *read.csv*.

```
> Mzdy <- read.table("Mzdy.txt", header = T) # dáta oddelené medzerou
```

Ak chceme načítať údaje s inými názvami stĺpcov (slovenskými), ako sú v súbore, stačí preskočiť prvý riadok s názvami cez argument *skip*, zadať *FALSE* pre argument *header*. To spôsobí načítanie s prednastavenými názvami V1 až V5, ktoré zmeníme ďalším príkazom *colnames*.

```
> Mzdy <- read.table("Mzdy.txt", header = F, skip = 1)
> colnames(Mzdy) <- c("mzda", "vzdelanie", "prax", "ženy", "vek")
```

Na import dát z takýchto typov súborov (iba s dátami bez inej pridanej informácie) môžeme použiť aj nástroj *Import Dataset* programu *RStudio*, ktorý sa spustí v okne so záložkou *Environment* a je intuitívny.



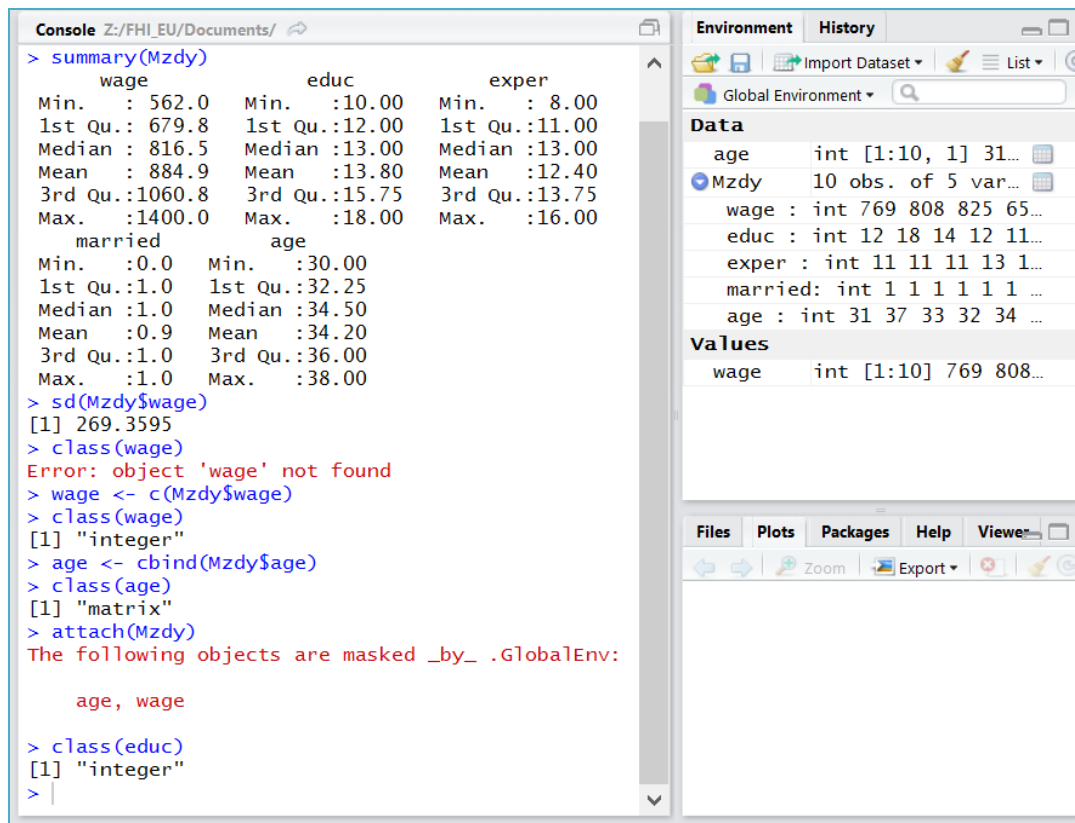
Obrázok 7: Import údajov cez nástroj *Import Dataset* programu *RStudio*

Importované údaje sú nahraté v relácii systému R ako špeciálny typ (potvrdí to príkaz *class(Mzdy)*), ktorý nazývame dátový rámec (*data frame*). Predstavuje zoznam vektorov s atribútmi, ako sú názvy stĺpcov. Prehľad o základných charakteristikách všetkých stĺpcov v dátovom rámci získame funkciou *summary(Mzdy)*. Ak nás zaujímajú charakteristiky iba jednotlivých premenných, môžeme využívať zabudované funkcie s odkazom na premennú v dátovom rámci cez znak \$, napr. v tvare *sd(Mzdy\$wage)*, ktorý zobrazí štandardnú odchýlku iba pre hodnoty v stĺpci *wage*.

```
> summary(Mzdy) # základné charakteristiky stĺpcov v dátovom rámci Mzdy
> sd(Mzdy$wage) # štand. odchýlka hodnôt wage v dátovom rámci Mzdy
```

Aby sme nemuseli vypisovať takéto dlhé reťazce, môžeme ich priradiť ku objektom s kratším názvom príkazmi *c* alebo *cbind* vysvetľovanými skôr alebo príkazom *attach* s názvom dátového rámca ako argumentom. Príkaz *attach* umožní prístup priamo k stĺpcom dátového rámca, Treba si dať pozor, ak máme nahratých viacero množín údajov, či neobsahujú premennú s rovnakým názvom.

```
> wage <- c(Mzdy$wage) # priradí stĺpec z dátového rámca do objektu wage
> age <- cbind(Mzdy$age) # priradí stĺpec z dátového rámca do vektora age
> attach(Mzdy) # základné charakteristiky stĺpcov v dátovom rámci Mzdy
```



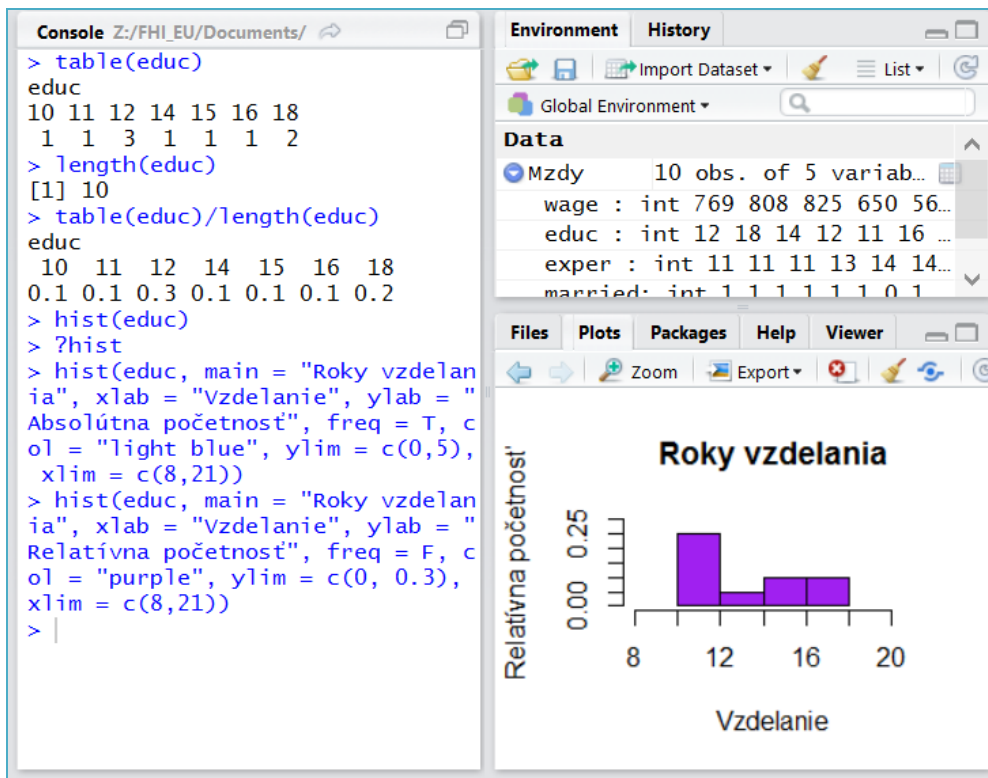
Obrázok 8: Prístup k importovaným údajom v dátových rámcoch

8 HISTOGRAM, BODOVÝ A ČIAROVÝ GRAF

Okrem výpočtov slúži programovací systém R aj na zobrazovanie dát pomocou grafov. Dva základné príkazy, ktoré sa pri tom využívajú sú *hist* a *plot*. Príkaz *hist* je určený na zobrazenie diskretných údajov (početností) pomocou histogramu. Na histograme môžeme zobrazit absolútne alebo aj relatívne početnosti. Absolútne početnosti zobrazíme v tabuľke príkazom *table*. Príkazom *length* získame jej počet pozorovaní. Relatívne početnosti sú absolútne početnosti predelené celkovým počtom pozorovaní.

Zadanie samotného argumentu názvu premennej za príkazom *hist* vykreslí čiernobiely histogram. Ďalšie argumenty ako: *main* zapíše názov grafu; *xlab* a *ylab* prepíšu preddefinované názvy popisiek; *col* zmení farbu stĺpcov; logický argument *freq* rozhodne, či sa zobrazia absolútne alebo relatívne početnosti a *xlim* a *ylim* definujú rozsah hodnôt na osiach grafu.

- > *table(educ)* # zobrazí tabuľku absolútnej početnosti pre *educ*
- > *length(educ)* # zobrazí počet pozorovaní pre *educ*
- > *table(educ)/length(educ)* # zobrazí tabuľku relatívnej početnosti pre *educ*
- > *hist(educ)* # vykreslí čiernobiely histogram absolútnej početnosti pre *educ*
- > *hist(educ, main = "Roky vzdelania", xlab = "Vzdelanie", ylab = "Absolútna početnosť", freq = T, col = "light blue", ylim = c(0,5), xlim = c(8,21))*
- > *hist(educ, main = "Roky vzdelania", xlab = "Vzdelanie", ylab = "Relatívna početnosť", freq = F, col = "purple", ylim = c(0,0.3), xlim = c(8,21))*



Obrázok 9: Početnosti a ich zobrazenie v tabuľke a na histograme

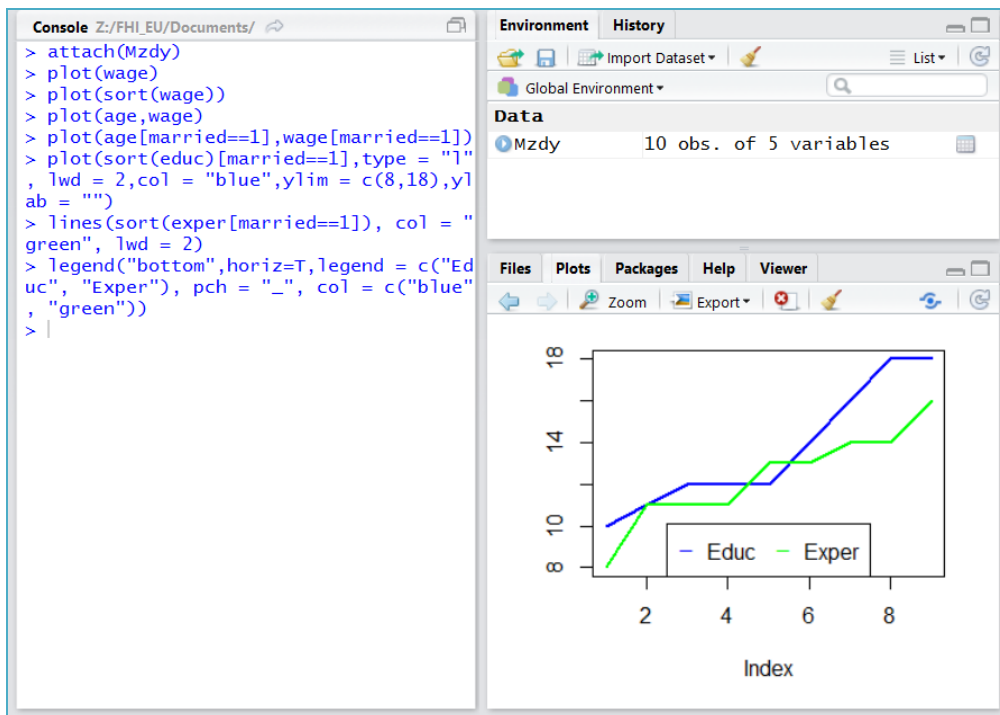
Grafy sa zobrazujú v programe RStudio v okne *Plots*, pričom vždy, keď zadáme nový príkaz zobrazenia grafu, zobrazuje sa iba posledný graf. Ak sa chceme vrátiť k predchádzajúcim grafom, použijeme šípky na paneli v okne *Plots* alebo príkazy v hlavnej ponuke *Plots* programu RStudio. Ak sa chceme pozerat' na zväčšený graf, využijeme *Zoom*, ktorý nám aktuálny graf zväčší a zobrazí v samostatnom okne. Grafy môžeme uložiť ako obrázky v rôznych formátoch alebo ako súbor typu *.pdf* prípadne ho skopírovať do schránky (*clipboard*). Pri každej akcii sa voľba musí presnejšie špecifikovať.

Okrem histogramov sa často využíva čiarový graf (*line*) alebo bodový graf XY (*scatter*). Zobrazíme ich príkazom *plot*. Ak za príkaz zadáme ako argument jednu premennú, zobrazí sa jej bodový priebeh v poradí údajov v premennej; ak dve premenné, zobrazí sa bodový graf dvojíc XY (*scatter*). Ďalšími argumentmi ako: *type* meníme typ grafu, potom *type = "l"* zmení graf na čiarový (*line*); *lwd* mení hrúbku čiary; *main* zadá názov grafu; *sub* zadá podnadpis grafu; *xlab* a *ylab* prepíšu preddefinované názvy popisiek; *col* zmení farbu stĺpcov. Výsledný graf môžeme doplniť príkazom *legend* o legendu alebo zobrazenie priebehu ďalšej premennej príkazom *lines*.

```

> plot(wage) # zobrazí bodový priebeh wage s indexom v poradí údajov
> plot(age, wage) # zobrazí bodový graf dvojíc wage a age
> plot(age[married==1], wage[married==1]) # zobrazí iba dvojice ženatých
> plot(sort(educ)[married==1], type="l", lwd=2, col="blue", ylim=c(8,18), ylab = "")
> lines(sort(exper[married==1]), col = "green", lwd=2) # pridá premennú
> legend("bottom", horiz = T, legend = c("Educ", "Exper"), pch = "_",
  col = c("blue", "green")) # pridá popis do grafu

```

Obrázok 10: Bodové a čiarové grafy premenných

9 INŠTALÁCIA PRÍDAVNÝCH BALÍKOV – PODPORA ANALÝZ

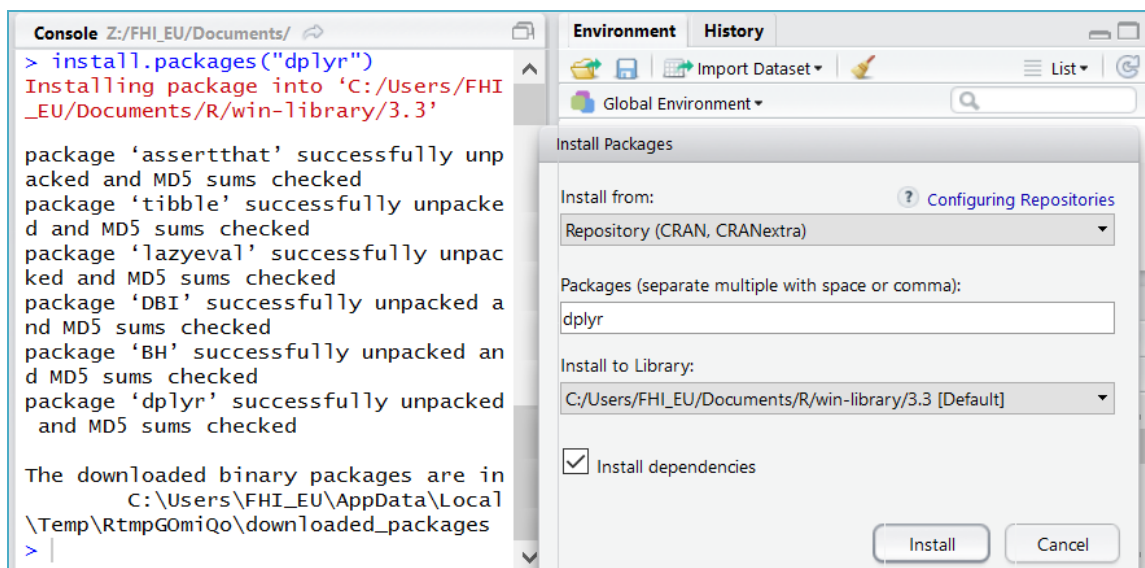
Výhodou programovacieho systému R je možnosť doinštalovať ľubovoľný prídavný balík obsahujúci naprogramované príkazy k rôznym oblastiam analýz alebo činnostiam pri práci s dátami. Informácie o prídavných balíkoch (*packages*), ktorých je vyše 9000, nájdeme na stránke https://cran.r-project.org/web/packages/available_packages_by_name.html.

V počítačoch s priamym pripojením na internet sa zvolený balík nainštaluje príkazom *install.packages* ktorého argumentom je jeho názov alebo v programe RStudio cez ponuku *Tools>Install_Packages*, kde sa názov balíka (alebo viacerých oddelených čiarkou) zadá do prostredného riadku. Používateľ môže v ponuke *Tools>Global_Options* v časti *Packages* zmeniť nastavený server (krajinu), odkiaľ sa balíky sťahujú.

Pri priamom používaní systému R nasleduje otázka voľby servera vždy po výbere *Install_package(s)* z ponuky *Packages*. Počas inštalácie balíka by mali byť zobrazované správy o jej priebehu. Po inštalácii sa balík aktivuje na použitie príkazom *library* alebo zaškrtnutím políčka pri názve balíka v okne *Packages* v programe RStudio. Pri práci s novou reláciou (napr. pri opätovnom spustení systému na ďalší deň), ak chceme využiť príkazy balíka, musíme ho nanovo aktivovať.

Ako príklad uvedieme príkazy na nainštalovanie balíka *dplyr*.

```
> install.packages("dplyr") # stiahne a nainštaluje prídavný balík dplyr
> library("dplyr") # aktivácia prídavného balíka dplyr pre danú reláciu
> help(package=dplyr) # zobrazí pomocníka k prídavnému balíku dplyr
```



Obrázok 11: Spustenie a priebeh inštalácie prídavného balíka (package)

Každý používateľ si môže vytvoriť vlastný prídavný balík, pomocou ktorého sa môže podeliť o svoj kód. Podrobný popis tvorby balíka nájdeme v knihe Wickhama (2015). Prídavné balíčky, ktoré boli vytvorené na podporu rôznych typov analýz (econometrics, bayesian inference) sú zhromaždené na internetových stránkach <https://cran.r-project.org/web/views/>.

Použitá literatúra

1. ALBERT, J., RIZZO, M. 2012. *R by Example*. New York: Springer, 2012.
2. BRUNSDON, C., COMBER, L. 2015. *An Introduction to R for Spatial Analysis and Mapping*. London: SAGE Publications, 2015.
3. CRAWLEY, M. J. 2013. *The R Book, 2nd. ed.* Chichester: John Wiley & Sons, 2013.
4. DAYAL, V. 2015. *An Introduction to R for Quantitative Economics Graphing, Simulating and Computing*. New Delhi: Springer, 2015.
5. EKSTRØM, C. T. 2012. *The R Primer*. Boca Raton: Taylor & Francis Group, 2012.
6. GERBING, D. W. 2014. *R Data Analysis without Programming*. Routledge: Taylor & Francis, 2014.
7. RODRIGUES, B. 2014. *Introduction to Programming Econometrics with R*. Dostupné na: <http://qe4policy.ec.unipi.it/wp-content/uploads/2015/10/Introduction-to-programming-Econometrics-with-R-Draft.pdf> [20.10.2016]
8. VENABLES, W. N., SMITH, D. M., the R Core Team. 2016. *An Introduction to R Notes on R: A Programming Environment for Data Analysis and Graphics*. Dostupné na: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf> [31.10.2016]
9. WICKHAM, H. 2015. *R Packages Organize, Test, Document, and Share Your Code*. Sebastopol: O'Reilly Media, 2015.

Kontaktné údaje

doc. Ing. Martin Lukáčik, PhD., Ing. Adriana Lukáčiková, PhD., doc. Ing. Karol Szomolányi, PhD.
 Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky
 Dolnozemska 1/b, 852 35 Bratislava
 Tel: (421 2) 67 295 822, (421 2) 67 295 829, (421 2) 67 295 822
 email: lukacik@euba.sk, szomolan@euba.sk, istvanik@euba.sk

MODEL TVORBY TURNUSOV V AUTOBUSOVEJ DOPRAVE S REŠPEKTOVANÍM BEZPEČNOSTNEJ PRESTÁVKY

DESIGN OF ROTATIONS IN PUBLIC BUS TRANSPORT WITH REGULAR BREAKS

Tomáš Majer, Stanislav Palúch

Abstrakt

V príspevku prezentujeme matematický model tvorby turnusov vo verejnej autobusovej doprave. Okrem klasických optimalizačných kritérií, akými sú minimalizácia počtu potrebných autobusov a minimalizácia dĺžky prázdnych prejazdov medzi spojmi sa zameriame na splnenie dodatočných podmienok kladených na pracovný čas vodičov pravidelnej dopravy, konkrétne na správne plánovanie bezpečnostnej prestávky po štyroch hodinách vedenia vozidla. Model experimentálne overíme na reálnych údajoch mestskej hromadnej dopravy v niektorom okresnom meste na Slovensku.

Kľúčové slová: turnusy vozidiel, zmiešané lineárne programovanie, bezpečnostná prestávka

Abstract

The paper presents a mathematical model of design of rotations in public bus transport. Besides classical optimization criteria such as minimizing the number of required buses and minimization of empty drives between journeys we focus on the fulfilment of additional conditions imposed on driver's working time, namely the proper planning of regular break after four hours of driving. Proposed model is experimentally verified on real data of urban public transport in a district town in Slovakia.

Keywords: vehicle scheduling, mixed linear programming, regular break

1 ÚVOD

Do osemdesiatych rokov minulého storočia sa plánovanie turnusov robilo ručne. S rozvojom výpočtovej techniky sa rozvíjajú aj metódy počítačovej tvorby turnusov – príklady niektorých prístupov rozvrhovania vozidiel vidieť v [1] prehľad rozvrhovania osádok je v [7].

Podmienky vo vtedajšom Československu sa výrazne líšili od podmienok vo vyspelých západných štátov, preto bolo treba vyvíjať vlastné optimalizačné metódy. Od roku 1980 sa v rámci štátnej výskumnej úlohy R-04-514-004 „Optimalizácia osobnej dopravy ČSAD“ na VÚD v Žiline rozvíjajú optimalizačné metódy tvorby turnusov, ktoré v r.1983 prvý úspešný výpočet turnusov znižujúci aj počet autobusov aj počet prázdnych najazdených km oproti skutočnému stavu. Po roku 1989 časť riešiteľského kolektívu prešla z VÚD na Žilinskú univerzitu, kde rozvoj rozvrhovacej problematiky pokračuje dodnes. Niektoré špeciálne výsledky možno nájsť v [3], [5], [6]. Súčasný stav tejto problematiky zachycuje práca [4].

Jednou zo špecifických podmienok kladených na turnusy vozidiel na Slovensku je, že autobus riadi len jeden alebo sa pri riadení striedajú dvaja vodiči. To znamená, že okrem bežných požiadaviek na presun autobusu v čase a priestore (autobus nemôže byť na dvoch miestach naraz), musia turnusy spĺňať aj požiadavky kladené na pracovný čas vodičov pravidelnej autobusovej dopravy. Tieto požiadavky sú definované v zákone 121/2004 o pracovnom čase

a dobe odpočinku v doprave a o zmene a doplnení niektorých zákonov. Z nich sa tvorby turnusov týka najmä bezpečnostná prestávka a prestávka na odpočinok a jedenie. V našom príspevku sa budeme zaoberať prvou menovanou.

2 MATEMATICKÝ MODEL

Spojom nazveme usporiadanú štvoricu $(cod_k, cpr_k, mod_k, mpr_k)$, kde cod_k je čas odchodu z prvej zastávky mod_k (miesto odchodu) a cpr_k je čas príchodu do poslednej zastávky mpr_k (miesto príchodu) spoja k .

Jednotlivé elementy matice $M = \{m(u, v)\}$ predstavujú čas potrebný na prázdny prejazd autobusu zo zastávky u do zastávky v .

Spoj j môže byť obsluhovaný tým istým autobusom po spoji i , ak platí:

$$cod_j \geq cpr_i + m(mpr_i, mod_j). \quad (1)$$

Inak povedané, ak sa autobus po ukončení obsluhy spoja i na zastávke mpr_i stihne presunúť na zastávku mod_j tak rýchlo, aby stihol začiatok obsluhy spoja j . Budeme tomu hovoriť, že spoj j môže nasledovať za spojom i a značiť $i < j$.

Turnus vozidla nazveme takú postupnosť spojov i_1, i_2, \dots, i_r , pre ktorú platí $i_1 < i_2 < \dots < i_r$, čiže postupnosť takých spojov, ktoré môžu byť obslužené jedným autobusom. Turnus predstavuje rozvrh práce pre jeden autobus na celý deň.

2.1 Minimalizácia počtu potrebných autobusov

Cieľom je použiť na obsluženie všetkých spojov čo najmenší počet autobusov, t.j. rozdeliť množinu spojov S do čo najmenšieho počtu turnusov tak, aby každý spoj patril práve do jedného turnusu. Takýto rozvrh budeme nazývať turnusový plán.

Reláciu následnosti spojov $<$ môžeme reprezentovať pomocou digrafu $\vec{G} = (S, E)$, kde $E = \{(i, j) | i \in S, j \in S, i < j\}$.

Nech množina spojov $S = \{1, 2, 3, \dots, n\}$. Binárna premenná x_{ij} bude vyjadrovať, či spoj j bude ($x_{ij} = 1$) alebo nebude ($x_{ij} = 0$) obsluhovaný bezprostredne za spojom i jedným autobusom v jednom turnuse.

Ak by boli všetky premenné $x_{ij} = 0$, potom by na obsluženie všetkých spojov bolo potrebných $|S| = n$ autobusov. Lebo každý spoj by bol vo svojom turnuse a všetky turnusy by obsahovali práve jeden spoj. Ak by práve jedna premenná x_{ij} nadobudla hodnotu 1, potom by bolo potrebných $n - 1$ autobusov, lebo práve jeden turnus by obsahoval dva spoje. Každá premenná x_{ij} , ktorá nadobudne hodnotu 1, ušetrí jeden autobus. Počet potrebných autobusov potom vyjadruje vzťah:

$$n - \sum_{(i,j) \in E} x_{ij}. \quad (2)$$

Minimalizovať počet autobusov znamená maximalizovať sumu $\sum_{(i,j) \in E} x_{ij}$. Matematický model na riešenie problému s minimálnym počtom turnusov by teda mal tvar:

$$\text{Maximalizuj} \quad \sum_{ij, (i,j) \in E} x_{ij} \quad (3)$$

za podmienok

$$\sum_{i,(i,j) \in E} x_{ij} \leq 1 \text{ pre } j = 1..n \quad (4)$$

$$\sum_{j,(i,j) \in E} x_{ij} \leq 1 \text{ pre } i = 1..n \quad (5)$$

$$x_{ij} \in \{0,1\} \text{ pre všetky } i, j = 1..n \quad (6)$$

Matematický model (3) až (6) predstavuje model prirad'ovacieho problému, ktorý bude dávať celočíselné riešenie aj bez podmienky (6), ktorá môže byť nahradená podmienkou:

$$x_{ij} \geq 0 \text{ pre všetky } i, j = 1..n. \quad (7)$$

Úloha navrhnúť turnusový plán s minimálnym počtom autobusov bez ďalších požiadaviek je teda riešiteľná v polynomiálnom čase.

2.2 Minimalizácia dĺžky prázdnych prejazdov

Vyriešením modelu (3)-(6) sme zistili potrebný počet autobusov, označme ho symbolom T . Ak zachovanie tohto počtu vynútime pridaním štruktúrálnej podmienky, môžeme účelovú funkciu zmeniť a minimalizovať súčet prázdnych prejazdov. Model by mal potom tvar:

Minimalizuj

$$\sum_{ij,(i,j) \in E} x_{ij} m(mpr_i, mod_j) \quad (8)$$

za podmienok

$$\sum_{i,(i,j) \in E} x_{ij} \leq 1 \text{ pre } j = 1..n \quad (9)$$

$$\sum_{j,(i,j) \in E} x_{ij} \leq 1 \text{ pre } i = 1..n \quad (10)$$

$$\sum_{ij,(i,j) \in E} x_{ij} = |S| - T \quad (11)$$

$$x_{ij} \in \{0,1\} \text{ pre všetky } i, j = 1..n \quad (12)$$

Tento model však nezohľadňuje prístavné a odstavné jazdy z/do garáže a minimalizuje len prejazdy medzi spojmi. Označme symbolom D depo (garáž), ktoré bude pre naše potreby špeciálna zastávka. Element $m(D, mod_i)$ matice prejazdov M predstavuje dĺžku prístavnej jazdy (jazdy z garáže na miesto odchodu spoja i) a element $m(mpr_j, D)$ matice prejazdov M predstavuje dĺžku odstavnej jazdy (jazdu z miesta príchodu spoja j do garáže).

Zaved'me binárne premenné u_j a v_i , ktoré budú nadobúdať hodnotu 1 práve vtedy, keď spoj j bude prvým resp. spoj i posledným spojom turnusu. Potom na minimalizáciu prázdnych prejazdov vrátane prístavnej a odstavnej jazdy môžeme použiť model:

Minimalizuj

$$\sum_{ij,(i,j) \in E} x_{ij} m(mpr_i, mod_j) + \sum_{j \in S} u_j m(D, mod_j) + \sum_{i \in S} v_i m(mpr_i, D) \quad (13)$$

$$\text{za podmienok} \quad u_j + \sum_{i,(i,j) \in E} x_{ij} = 1 \text{ pre } j = 1..n \quad (14)$$

$$v_i + \sum_{j,(i,j) \in E} x_{ij} = 1 \text{ pre } i = 1..n \quad (15)$$

$$\sum_{ij,(i,j) \in E} x_{ij} = |S| - T \quad (16)$$

$$x_{ij}, u_j, v_i \in \{0,1\} \text{ pre všetky } i, j = 1..n \quad (17)$$

2.3 Kontrola bezpečnostnej prestávky

Bezpečnostná prestávka musí nasledovať najneskôr po 4 hodinách vedenia vozidla a mala by trvať minimálne 30 minút. Bezpečnostnú prestávku možno rozdeliť aj na viac častí, každá časť ale musí trvať minimálne 10 minút. Prerušenie vedenia vozidla na menej ako 10 minút sa započítava do času vedenia vozidla.

Bezpečnostnú prestávku budeme kontrolovať v už hotových turnusoch. Keď spočítame čas vedenia vozidla v ľubovoľnom časovom okne dĺžky 270 minút (4,5 hodiny), čas vedenia vozidla nesmie presiahnuť 240 minút. Príklad časti turnusu, kde došlo k porušeniu ustanovenia o bezpečnostnej prestávke je znázornený v tabuľke Tab. 1.

Linka	Spoj	Miesto odchodu	Čas odchodu	Miesto príchodu	Čas odchodu	Prejazd pred spojom	Prejazd po spoji	Čas státia
27	25	59	13:27	70	14:00	0	0	5
27	28	70	14:05	59	14:38	0	0	21
7	51	59	14:59	39	15:25	0	0	4
7	52	39	15:29	59	15:54	0	0	0
12	17	59	15:54	24	16:21	0	0	5
12	18	24	16:26	56	16:50	0	2	7
12	19	18	16:59	24	17:21	0	0	5
12	20	24	17:26	59	17:54	0	0	0

Tab.1: Ukážka časti turnusu s porušeným pravidlom o bezpečnostnej prestávke

V 270 minútovom časovom okne od 13:27 do 17:57 je jedinou prestávkou v dĺžke viac ako 10 minút prestávka v čase od 14:38 do 14:59 nasledujúca po spoji 28 na linke 27. Táto prestávka má len 21 minút a teda čas vedenia vozidla v tomto časovom okne je 249 minút, t.j. viac ako 240 minút.

Ak chceme zakázať výskyt takejto časti turnusu, musíme do modelu (13) až (17) pridať štruktúrnu podmienku, ktorá zakáže výskyt všetkých siedmych spojov za sebou v jednom turnuse, t.j. so siedmych premenných x_{ij} môže nadobúdať hodnotu 1 maximálne 6 z nich:

$$x_{27/25,27/28} + x_{27/28,7/51} + x_{7/51,7/52} + x_{7/52,12/17} + x_{12/17,12/18} + x_{12/18,12/19} + x_{12/19,12/20} \leq 6 \quad (18)$$

Podobnú podmienku musíme pridať pre všetky porušenia bezpečnostnej prestávky. Po výpočte modelu (13) až (17) spolu s niekoľkými podmienkami typu (18) sa môže stať, že opäť dôjde k porušeniu ustanovenia o bezpečnostnej prestávke, ale už zreťazením iných spojov. Ak sa tak stane, pridáme do modelu ďalšie podmienky a model znovu vypočítame. Postup opakujeme až dovtedy, kým budú všetky mať turnusy bezpečnostnú prestávku na správnom mieste. Vzhľadom na znižovanie množiny prípustných riešení sa tiež môže stať, že model nebude mať prípustné riešenie. To by znamenalo, že treba zvýšiť počet turnusov, lebo nie je možné urobiť turnusový plán s minimálnym počtom autobusov a zároveň dodržať ustanovenie o bezpečnostnej prestávke.

3 EXPERIMENT

Navrhnuté matematické modely sme použili na určenie turnusov pre spoje mestskej hromadnej dopravy premávajúce pred pár rokmi v súmestí Martin – Vrútky.

Ide o 723 spojov premávajúcich na 17 linkách cez pracovný deň, garáž je v číselníku zastávok reprezentovaná číslom 59.

V prvom kroku sme vyriešením modelu (3) až (6) zistili, že turnusový plán bude pozostávať z 39 turnusov, t.j. $T = 39$. Následne sme tento počet turnusov použili v modeli (13) až (17). Zostavenie týchto modelov skriptom v jazyku Python a ich vyriešenie pritom trvalo cca 62 resp. 66 sekúnd.

Potom sme analýzou riešenia modelu (13) až (17) zistili niekoľko porušení ustanovenia o bezpečnostnej prestávke a pre každé porušenie sme do modelu pridali podmienku typu podmienky (18), samozrejme pre správne spoje z analyzovaného turnusu. Časy opakovaného riešenia modelu s postupne pridávanými podmienkami uvádzame v tabuľke Tab. 2.

Iterácia	Čas zostavenia resp. úpravy modelu	Čas výpočtu modelu	Počet porušení bezp.prestávky
1	56,43 s	7,38 s	7
2	0,34 s	7,59 s	2
3	0,33 s	7,31 s	9
4	0,34 s	8,17 s	6
5	0,33 s	8,59 s	8
6	0,34 s	7,82 s	5
7	0,34 s	8,69 s	0

Tab.2: Časy výpočtu jednotlivých iterácií pri odstraňovaní porušení ustanovenia o bezpečnostnej prestávke

Celkový čas opakovaného riešenia bol cca 112 sekúnd a všetkých 39 turnusov po siedmej iterácii spĺňalo ustanovenie o bezpečnostnej prestávke.

4 ZÁVER

V príspevku sme definovali problém tvorby turnusov a formulovali MILP matematické modely na určenie minimálneho počtu potrebných autobusov a navrhovanie turnusov

s minimálnou dĺžkou prázdnych prejazdov. Tiež sme navrhli postup, ako opakovaným výpočtom modelu s pridanými podmienkami zakazujúcimi porušovanie pravidla o bezpečnostnej prestávke pre konkrétnu postupnosť spojov dosiahnuť vytvorenie turnusov s minimálnou dĺžkou prázdnych prejazdov a rešpektovaním ustanovenia o bezpečnostnej prestávke podľa príslušného zákona.

PodĎakovanie

Tento príspevok vznikol vďaka podpore grantu APVV-14-0658 "Optimalizácia mestskej a regionálnej verejnej dopravy".

Použitá literatúra

1. Ceder, A.: Public-transit vehicle schedules using a minimum crew-cost approach, Total Logistic Management vol. 4, pp 21-42, 2011.
2. Gurobi Optimization Inc.: Gurobi Optimizer Reference Manual, dostupné na <http://www.gurobi.com/documentation/6.5/refman.pdf>, 2015.
3. Palúch, S.: Two approaches to vehicle and crew scheduling in urban and regional bus transport, Proceedings of international conference Quantitative Methods in Economics 2008, High Tatras, Slovakia, pp. 212-218, 2008.
4. Palúch, S.: Vehicle and crew scheduling problem in regular personal bus transport - the state of art in Slovakia, Proceedings of 16th international conference on transport science ICTS 2013, Portorož, Slovenija, pp. 297-304, 2013.
5. Peško, Š.: The minimum fleet size problem for flexible bus scheduling, Studies of the Faculty of management science and informatics University of Žilina vol. 9(1), pp. 75-81, 2001.
6. Peško, Š.: Flexible bus scheduling with overlapping trips, Proceedings of international conference Quantitative Methods in Economics 2008, High Tatras, Slovakia, pp. 225-230, 2008.
7. Wren, A., Rousseau, J.M.: Bus driver scheduling - an overview, Computer-Aided Transit Scheduling, Springer Verlag, 1995.

Kontaktné údaje

Ing. Tomáš Majer, PhD.

Žilinská univerzita v Žiline, Fakulta riadenia a informatiky

Univerzitná 8215/1, 010 26 Žilina

Tel: (421 41) 513 4284

email: tomas.majer@fri.uniza.sk

doc. RNDr. Stanislav Palúch, CSc.

Žilinská univerzita v Žiline, Fakulta riadenia a informatiky

Univerzitná 8215/1, 010 26 Žilina

Tel: (421 41) 513 4250

email: stanislav.paluch@fri.uniza.sk

MODELOVANIE PRÍJMOV DOMÁCNOSTÍ V CGE MODELI

MODELING THE HOUSEHOLDS' INCOME IN THE CGE MODEL

Veronika Miťková, Eduard Hozlár

Abstrakt

Článok sa venuje odhadu parametrov podielu vo funkciách príjmov domácností v modeli všeobecnej ekonomickej rovnováhy. Použitou metódou je ekonometrický odhad, sledovaným obdobím sú roky 1995 až 2014. Regionálne členenie je na Bratislavský kraj a ostatné kraje Slovenskej republiky, s dezagregáciou verejného sektora a ostatných sektorov a výrobných faktorov na prácu a kapitál. Splnenie podmienok klasického ekonometrického modelu vedie k deformácii pôvodných funkčných vzťahov definovaných modelom a bude potrebné voliť iné metódy odhadu tak, aby boli dodržané.

KLúčové slová: príjem domácnosti, CGE model, ekonometrický odhad

Abstract

The paper deals with the estimation of the share parameters of households' income functions in the computable general equilibrium model. The methodology of econometric estimation is used, the time range of 1995 till 2014. Regional disaggregation to Bratislava region and rest of the regions of the Slovak Republic, commodity disaggregation to the public sector and other sectors, and labor and capital factors. Fulfillment of the classic econometric model leads to the deformation of original functional forms defined by the model and there is a need to find another methods of estimation to satisfy them.

Keywords: household's income, CGE model, econometric estimation

1 ÚVOD

V článku sa venujeme ekonometrickým odhadom parametrov funkcií príjmov domácností v modeli všeobecnej ekonomickej rovnováhy (CGE – Computable General Equilibrium model). Na odhady hodnôt parametrov sú tradične používané kalibračné metódy, ktoré vychádzajú z poznania stavu ekonomiky k určitému času. Ich výhodou je možnosť získať hodnoty parametrov modelu z databázy, ktorá nevyžaduje dlhé časové rady, ako je tomu v ekonometrickom modelovaní. Využívajú bodový odhad, ktorý je však silne ovplyvnený voľbou časového obdobia, za ktoré konštruujeme databázu údajov, čo môže viesť k skresleným odhadom hodnôt napr. elasticít, parametrov posunu a podielu. Príkladom môžu byť turbulentné obdobia predkrízových období, roky prasknutia realitných, internetových či iných bublín, ktoré majú následne dosah na hodnoty makroekonomických agregátov, a tým aj hodnôt parametrov modelu. Následná kalibrácia hodnôt parametrov CGE modelu potom vedie k takým hodnotám, ktoré sú síce reprezentantom príslušného obdobia, nie je však vhodné využívať ich na experimenty s modelom v iných obdobiach. Ekonometrický odhad túto slabinu eliminuje, vyžaduje však dáta dlhých časových období, ktoré nie sú vždy dostupné. Kritikou kalibrácie a metodológiou ekonometrických odhadov sa zaoberajú napr. Jorgenson (1984), Lau (1984), ich práce boli nasledované McKitrick (1988) a Diewert a Lawrence (1994). V nedávnej minulosti boli vyvinuté techniky kombinujúce rôzne prístupy

k odhadom parametrov CGE modelov, ktoré využívajú všetky dostupné údaje, pozri napr. Arndt et al. (2002).

2 MODEL VŠEOBECNEJ EKONOMICKEJ ROVNOVÁHY

Cieľom tohto článku je odhad parametrov funkcií príjmov domácností v CGE modeli pomocou ekonometrického odhadu založenom na časových radoch pre Slovenskú republiku za obdobie 1995q1 až 2014q4. Údaje nie sú sezónne očistené, zámerne pracujeme s neupravenými dátami. Časové rady sú dezagregované z regionálneho hľadiska na sektor Bratislavského kraja (BA) a zvyšku Slovenska (SK), z hľadiska komodít na agregovaný verejný sektor (PS), ktorý zahŕňa verejnú správu, obranu, zdravotníctvo, školstvo a sociálne služby a na agregovaný sektor ostatných komodít (NS). Výrobné faktory predpokladáme prácu (LAB) a kapitál (CAP). Štruktúru modelu vyvinuli na International Food Policy Research Institute, autorom je Löfgren (2003), tento bol však čiastočne modifikovaný tak, aby vyhovoval podmienkam modelovanej ekonomiky, pozri Miťková (2015). Model pozostáva celkovo zo 17 rovníc: produkčné funkcie, dopyt po výrobných faktoroch, príjem výrobných faktorov, príjem domácností, dopyt domácností, rovnovážne podmienky ponuky a dopytu výrobného faktora a funkcia normalizácie cien.

Predpokladáme, že príjmy domácnosti závisia od priemerných cien kapitálu a práce a od dopytu po týchto výrobných faktoroch tak, ako to uvádzajú funkcie 1 – 4:

$$YH_{BA,CAP,t} = shry_{BA,CAP} \cdot WF_{CAP,t} \cdot (QF_{CAP,PS,t} + QF_{CAP,NS,t}) + e_{t1} \quad (1)$$

$$YH_{BA,LAB,t} = shry_{BA,LAB} \cdot WF_{LAB,t} \cdot (QF_{LAB,PS,t} + QF_{LAB,NS,t}) + e_{t2} \quad (2)$$

$$YH_{SK,CAP,t} = shry_{SK,CAP} \cdot WF_{CAP,t} \cdot (QF_{CAP,PS,t} + QF_{CAP,NS,t}) + e_{t3} \quad (3)$$

$$YH_{SK,LAB,t} = shry_{SK,LAB} \cdot WF_{LAB,t} \cdot (QF_{LAB,PS,t} + QF_{LAB,NS,t}) + e_{t4} \quad (4)$$

kde: YH_{hf} transfer príjmu domácnosti $h \in \{BA, SK\}$ od výrobného faktora $f \in \{CAP, LAB\}$

,
 WF_f priemerná cena výrobného faktora $f \in \{CAP, LAB\}$,

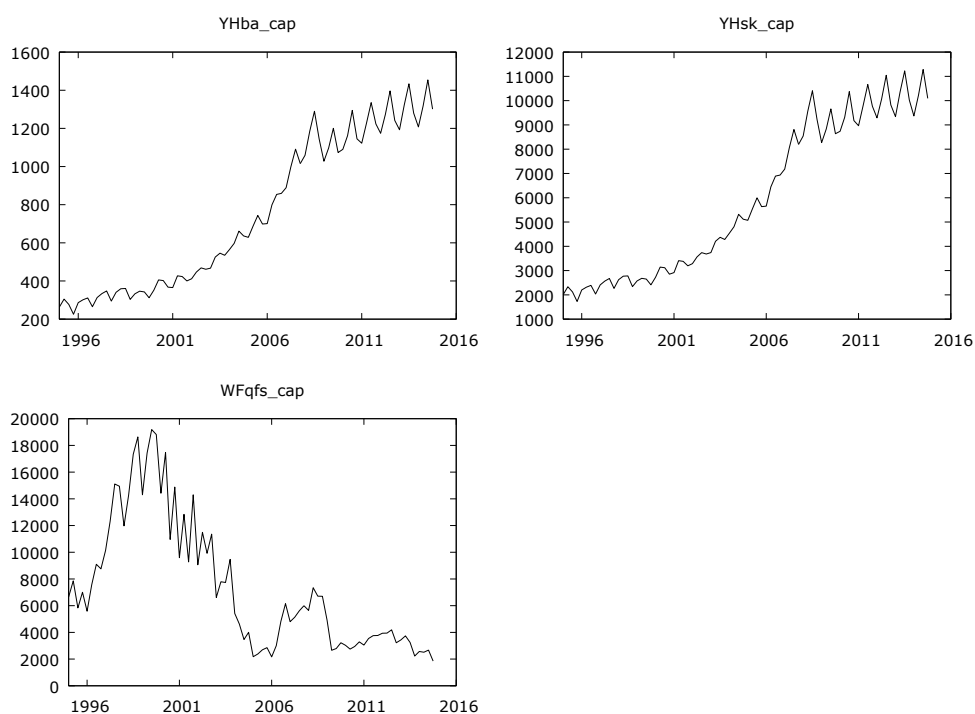
QF_{fa} požadované množstvo výrobného faktora $f \in \{CAP, LAB\}$ aktivitou $a \in \{PS, NS\}$,

$shry_{hf}$ parameter podielu príjmu domácnosti $h \in \{BA, SK\}$ z vlastníctva výrobného faktora $f \in \{CAP, LAB\}$.

V týchto rovniciach je závislou premennou príjem domácností z vlastníctva výrobného faktora a vysvetľujúcou premennou je násobok ceny a dopytu po výrobnom faktore (čo je súčet dopytov dvoch domácností), ktorý vyjadruje hodnotu dopytu po výrobnom faktore (pracovné označenie týchto premenných je $WFQFS_{cap}$ a $WFQFS_{lab}$). Odhadujeme hodnotu parametra podielu príjmu domácností z vlastníctva výrobných faktorov.

2.1 Príjem z kapitálu

Modelovanie príjmu z vlastníctva kapitálu (rovnice 1 a 3) sa sústreďuje prevažne na technickú stránku vzťahu, keďže časové rady príjmu domácností a hodnoty dopytu po kapitále majú opačné trendy s koreláciou - 0,71, ako to zobrazuje Obrázok 1. Časový rad vľavo hore zobrazuje príjem bratislavských domácností z vlastníctva kapitálu, rad vpravo hore zase príjem domácností západného, stredného a východného Slovenska z vlastníctva kapitálu. Dolný časový rad zobrazuje hodnotu dopytu po kapitáli. Údaje sú v nominálnych hodnotách, nie sú sezónne očistené.



Obrázok 1: Časový rad príjmov domácností (horné obrázky) a hodnoty dopytu po kapitále (dolný obrázok).

Počiatočný ekonometrický odhad rovnice (1) OLS metódou neprináša uspokojivý výsledok, ako sa to aj dalo očakávať na základe priebehu časových radov a korelačného koeficientu. Variabilita je v modeli vysvetlená len na 31,02%. So zreteľom na čo najmenšie zmeny vo funkčnej štruktúre modelu sme testovali rôzne modifikácie modelu a ako najvhodnejší sa javí model s pôvodnými vysvetľujúcimi premennými a pridaným časovým trendom. Dostávame vysoko štatisticky významné odhady parametrov, avšak s prítomnosťou autokorelácie a nie normálne rozdelenými rezíduami. Odhad Prais-Winstenovou iteratívnou technikou odstraňuje autokoreláciu, rezíduá majú normálne rozdelenie, avšak vysvetľujúce premenné nie sú štatisticky významné. Hodnoty odhadov použitím oboch metód sú veľmi blízke, s nižšími štandardnými odchýlkami v prospech OLS, takže akceptujeme výsledky tohto odhadu.

$$\widehat{Y}_{BA,CAP,t} = 0.0035 \cdot WF_{CAP,t}(QF_{CAP,PS,t} + QF_{CAP,NS,t}) + 17.6514 \cdot t \quad (5)$$

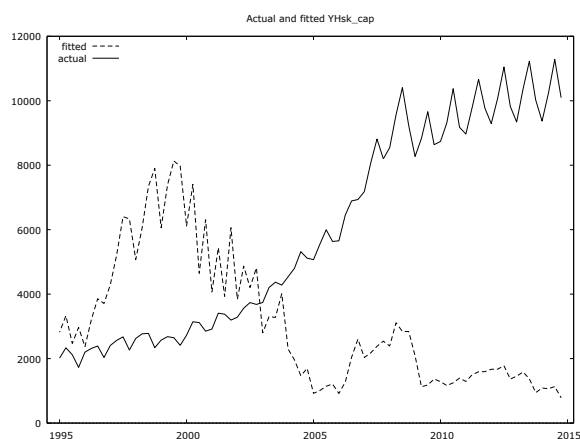
Model zachytáva 98,33% variability závislej premennej, takže akceptujeme porušenie predpokladu o normalite rezíduí. Tento problém môže byť odstránený pridaním sezónnych umelých premenných, ktoré však spôsobili nevýznamnosť pôvodných vysvetľujúcich

premenných a neboli teda ďalej uvažované. Sezónne diferencie neodstránili problém porušenia tohto predpokladu, navyše viedli k odlišnej modelovej štruktúre. Hodnota kapitálu má v oboch sektoroch len malý vplyv na transfer príjmu bratislavským domácnostiam, najvyšší je vplyv času indikujúci štvrtročný rast príjmu domácností o 17,65 mil. Eur.

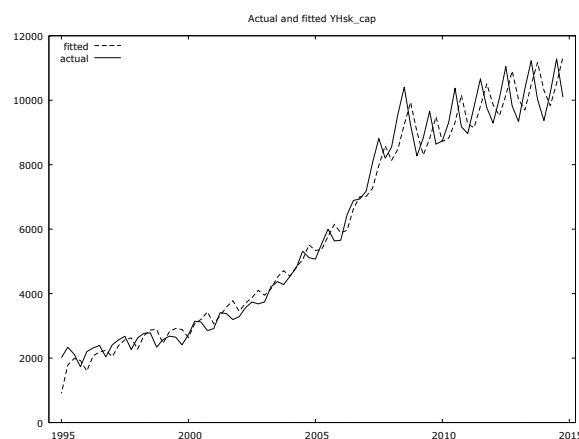
Obdobná je situácia v rovnici vzťahu príjmu mimobratisklavských domácností z vlastníctva kapitálu. Obrázok 2 ilustruje zlyhanie nielen OLS, rovnaký priebeh majú vyrovnané časové rady získané pomocou 2-SLS alebo FIML. Uvažovali sme znova pridanie časového trendu, Prais-Winstenova metóda (6) vedie k odhadu bez autokorelácie, s normálne distribuovanými rezíduami, avšak so štatisticky nevýznamnou pôvodnou vysvetľujúcou premennou. Príjmy slovenských domácností (mimo bratislavského kraja) z vlastníctva kapitálu sa štvrtročne zvyšujú o 138,89 mil. Eur.

$$\widehat{YH}_{SK,CAP,t} = 0.0243 \cdot WF_{CAP,t}(QF_{CAP,PS,t} + QF_{CAP,NS,t}) + 138.882 \cdot t \quad (6)$$

2.2



Obrázok 2: Časový rad príjmu Slovenských (mimobratisklavských) domácností z vlastníctva kapitálu – pôvodné a vyrovnané hodnoty pomocou OLS.

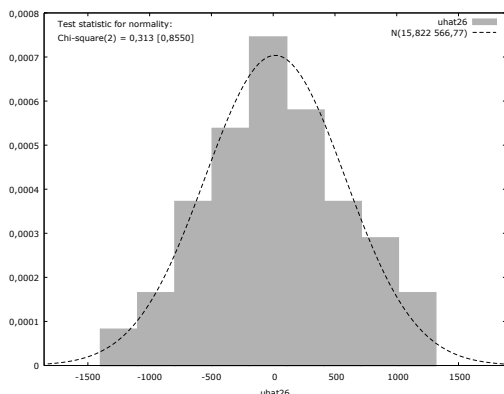


Obrázok 3: Časový rad príjmu Slovenských (mimobratisklavských) domácností z vlastníctva kapitálu – pôvodné a vyrovnané hodnoty pomocou Prais-Winsten s časovým trendom.

2.3 Príjem z práce

Modelovanie príjmov domácností z vlastníctva výrobného faktora práca je priamočiarejšie, nakoľko ich korelácia s hodnotou dopytu po tomto výrobnom faktore je vysoká a kladná: 0,99. Príjem bratislavských domácností v závislosti od hodnoty dopytu po práci vykazuje vysokú mieru vysvetlenej variability, avšak s prítomnosťou autokorelácie náhodnej zložky. Táto bola odstránená použitím Prais-Winstenovej iteratívnej metódy, ktorá dáva spomedzi všetkých odhadov korekcie autokorelácie najlepšie výsledky. Model zachytáva 97,89% variability príjmov, predpokladu normality rezíduí je v tomto prípade splnený – Obrázok 4.

$$\widehat{YH}_{BA,LAB,t} = 0.0397 \cdot WF_{LAB,t}(QF_{LAB,PS,t} + QF_{LAB,NS,t}) \quad (7)$$



Obrázok 4: Rozdelenie rezíduí pre model príjmov mimobratislavských domácností.

Parameter podielu príjmu slovenských domácností z vlastníctva práce bol odhadovaný pomocou metódy Hildreth-Lu, a to práve pre autokorelácie prvého rádu. Tento bol odstránený, pretrváva však problém normality rezíduí, ktorý v tomto odhade nie je splnený.

$$\hat{Y}_{SK,LAB,t} = 0.323 \cdot WF_{LAB,t} (QF_{LAB,PS,t} + QF_{LAB,NS,t}) \quad (8)$$

Parameter podielu príjmu z práce pre mimobratislavské domácnosti indikuje vysokú mieru príjmov z tohto výrobného faktora.

3 ZÁVER

Ekonometrické odhady korigované v zmysle naplnenia klasických predpokladov o náhodnej chybe často vedú k zmenám funkčných tvarov rovníc, a tým k deformácii pomerových parametrov. Pre porovnanie uvádzame sumarizáciu parametrov podielu pre príjem domácností z vlastníctva výrobných faktorov kapitál a práca: v Tabuľke 1 tak, ako ich uvádzame v tomto článku, v Tabuľke 2 bez odstránenia porúch.

Tabuľka 1: Odhady parametrov podielu so zachovaním funkčných tvarov rovníc

$shry_{hf}$	CAP	LAB	Σ
BA	0,0536	0,0648	0,1184
SK	0,4238	0,5143	0,9381
Σ	0,4774	0,5791	1,0565

Tabuľka 2: Odhady parametrov podielu s korekciou porúch klasického modelu

$shry_{hf}$	CAP	LAB	Σ
BA	0,0035	0,0397	0,0432
SK	0,0243	0,323	0,3473
Σ	0,0278	0,3627	0,3905

Odhady bez modifikácií funkčných tvarov rovníc sa približujú podmienke $\sum_h \sum_f shry_{hf} = 1$, ktorá je pre CGE model kľúčová. V ďalšom výskume sa preto sústreďme na také techniky odhadu parametrov modelu, ktoré umožnia zachovať túto podmienku.

3.1 Podpora výskumu

Výskum opísaný v tomto článku bol podporený Agentúrou pre vedu a výskum, projekt č. APVV-14-0020: Pareto Optimal Interactions Modeling in the Financial and Business Adaptive Processes a Vedeckou a edukačnou grantovou agentúrou, projekt č. VEGA 1/0340/16: “Modelovanie efektívnosti financovania a kooperácie sektorov a regiónov slovenskej ekonomiky.”

Použitá literatúra

1. ARNDT, C., S. ROBINSON, F. TARP, F. 2002. Parameter Estimation for a Computable General Equilibrium Model: A Maximum Entropy Approach. In: *Economic Modelling* 19. pp. 375-398.
2. DIEWERT, W.E., D. LAWRENCE 1994. The Marginal Costs of Taxation in New Zeland. Report prepared for the New Zeland Business Round Table by Swan Consultants Pty Ltd., Canberra.
3. JORGENSON, D.W. 1984. Econometric methods for applied general equilibrium analysis. In: Scarf, H.E. and J.B. Shoven (eds.) *Applied General Equilibrium Analysis*. Cambridge: Cambridge University Press.
4. LAU, L. 1984. Comments on Mansur and Whalley's Numerical Specification of Applied General Equilibrium Models. In: Scarf, H.E. and J. B. Shoven (eds.) *Applied General Equilibrium Analysis*. Cambridge: Cambridge University Press.
5. LÖFGREN, H. 2003. Exercises in General Equilibrium Modeling Using GAMS. Microcomputers in Policy Research 4a. Washington: International Food Policy Research Institute.
6. MCKITRICK, R.R. 1988. The Econometric Critique of Computable General Equilibrium Modeling: The Role of Functional Forms. *Economic Modelling* 15.
7. MIŤKOVÁ, V. 2015. Computable equilibrium model parameters calibration vs. estimation. In: Zborník zo zahraničnej vedeckej konferencie *Mathematical Methods in Economics 2015*, Plzeň: University of West Bohemia. ISBN 978-80-261-0539-8.

Kontaktné údaje

Ing. Veronika Miťková, PhD.

Ústav ekonómie, Fakulta sociálnych a ekonomických vied Univerzity Komenského v Bratislave

Mlynské luhy 4, 821 05 Bratislava

email: veronika.mitkova@fses.uniba.sk

doc. RNDr. Eduard Hozlár, CSc.

Ústav ekonómie, Fakulta sociálnych a ekonomických vied Univerzity Komenského v Bratislave

Mlynské luhy 4, 821 05 Bratislava

email: eduard.hozlar@fses.uniba.sk

MAKROEKONOMICKÝ VÝVOJ A IDENTIFIKÁCIA INVESTIČNÝCH PRÍLEŽITOSTÍ¹

MACROECONOMIC DEVELOPMENT AND INVESTMENT OPPORTUNITIES IDENTIFICATION

Vladimír Mlynarovič

Abstrakt

Článok prezentuje makroekonomický prístup k identifikácii investičných príležitostí medzi vládnymi dlhopismi. Je založený na konfrontácii rozpätí ich výnosov a úrovni a trendov vo vývoji makroekonomických charakteristík zodpovedajúcich krajín. Rozpätie definované ako rozdiel medzi výnosom do splatnosti vládneho dlhopisu v konkrétnej krajine a výnosom nemeckého vládneho dlhopisu s tou istou splatnosťou. Na makroekonomické zoradenie vybraných krajín sa využívajú historické a prognózované údaje publikované OECD. Vybrané kritéria ktoré merajú trendy a úrovne sú transformované na tzv. z-scores a metódy triedy Promethee sú použité na výsledné zoradenie.

KLúčové slová: makroekonomické kritériá, rozpätie výnosov (spreads), viackriteriálny prístup

Abstract

The paper presents macroeconomic approach for identification of investment opportunities among government bonds. It is based on a confrontation of their spreads and level and trends characteristic of the macroeconomic development of corresponding countries. The spread is defined as the difference between the yield to maturity of the specific country government bond and Germany government bond with the same maturity. The OECD historical and prognostic macroeconomic data are used for macroeconomic ranking of selected countries. Selected criteria that measures trends and levels are transformed into so called z-scores and application of Promethee methodology provides the final ranking.

Keywords: macroeconomic criteria, spreads, multiple criteria approach

1 ÚVOD

Existuje veľa aplikácii techník komplexného vyhodnocovania variantov, ktoré ako hlavný a jediný výsledok ponúkajú zoradenie skúmaných variantov alebo alternatív. V mnohých prípadoch dokonca chýba aj akákoľvek analýza senzitivnosti k možným zmenám vstupných údajov. V takýchto prípadoch môže vzniknúť otázka týkajúca sa užitočnosti takýchto výsledkov. V tomto príspevku síce tiež využívame takéto techniky na zoradenie krajín na báze vývoja ich makroekonomických charakteristík vývoja, ale tieto výsledky sú konfrontované s rozpätiami výnosov vládnych dlhopisov s cieľom identifikovať sľubné investičné príležitosti v tejto kategórii aktív

¹ 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."

2 MODEL

Predpokladajme, že poznáme hodnoty a_{ij}^t jednotlivých variantov (krajín) $i, i = 1, \dots, m$, podľa jednotlivých kritérií (makroekonomických charakteristík) $j, j = 1, \dots, n$, pre špecifikované časové obdobia (roky) $t, t = 1, \dots, T$. Pre tieto hodnoty definujeme dva typy makroekonomických kritérií, a to kritéria úrovne a kritéria trendu. Nech $[d_l, h_l], 1 \leq d_l \leq h_l \leq T$, je časové obdobie pre kritéria úrovne a $[d_r, h_r], 1 \leq d_r < d_l \leq h_l \leq h_r \leq T$, je časové obdobie pre kritéria trendu, potom

$$c_{ij}^l = \frac{\sum_{t=d_l}^{h_l} a_{ij}^t}{h_l - d_l + 1}$$

je hodnota úrovne pre kritérium $j, j = 1, \dots, n$, pre krajinu $i, i = 1, \dots, m$, a

$$c_{ij}^r = c_{ij}^l - \frac{\sum_{t=d_r}^{h_r} a_{ij}^t}{h_r - d_r + 1}$$

je hodnota trendu pre kritérium $j, j = 1, \dots, n$, pre krajinu $i, i = 1, \dots, m$.

Takže máme úlohu viackriteriálneho rozhodovania s m alternatívami and $2n$ kritériami a môžeme použiť niektoré so známych metód, (Mlynarovič, 1998) na zoradenie alternatív. V tejto aplikácii ale navrhujeme použiť tzv. z – scores namiesto priameho použitia hodnôt úrovni a trendov podľa kritérií. Výhodou takéhoto prístupu je skutočnosť že sa zohľadňujú nielen hodnoty kritérií ale aj ich variabilita. Nech μ_j^l a σ_j^l je stredná hodnota a štandardná odchýlka hodnôt $c_{ij}^l, i = 1, \dots, m$, a μ_j^r and σ_j^r je stredná hodnota a štandardná odchýlka hodnôt $c_{ij}^r, i = 1, \dots, m$, potom

$$s_{ij}^l = \frac{c_{ij}^l - \mu_j^l}{\sigma_j^l}, \quad i = 1, \dots, m, j = 1, \dots, n$$

sú hodnoty z – scores pre kritériá úrovne a

$$s_{ij}^r = \frac{c_{ij}^r - \mu_j^r}{\sigma_j^r}, \quad i = 1, \dots, m, j = 1, \dots, n$$

sú hodnoty z – scores pre kritériá trendu.

Makroekonomické zoradenie krajín možno zapísať ako úlohu komplexného vyhodnocovania variantov

$$\text{"max"} \quad \{s_i = (s_i^l, s_i^r) = (s_{i1}^l, \dots, s_{in}^l, s_{i1}^r, \dots, s_{in}^r) \quad i = 1, \dots, m\}$$

kde sa, bez straty všeobecnosti, predpokladá, že “čím viac, tým lepšie” pre všetky kritéria. Predpokladajme, že aplikácia metódy PROMETHEE II poskytne hodnoty čistých tokov $\Phi_i, i = 1, \dots, m$, pre ktoré

$$\Phi_1 > \Phi_2 > \dots > \Phi_i > \dots > \Phi_m$$

V terminológii metódy PROMETHEE to znamená, že krajina 1 je najlepšia a krajina m je najhoršia krajina.

Predpokladajme, že $r_i, i = 1, \dots, m$, výnos do splatnosti vládnych dlhopisov ponúkaný krajinou i . Potom vo svete ideálnych informácií by pre uvedené PROMETHEE výsledky malo platiť

$$r_1 < r_2 < \dots < r_i < \dots < r_m$$

pretože čím horšie je na tom krajina, tým vyššie výnosy musí ponúkať. V tejto situácii sa potenciálny investor musí rozhodnúť, aké veľké riziko je ochotný akceptovať. Prakticky ťažko možno očakávať takýto jednoznačný výsledok a možné kontradikcie medzi makroekonomickými výsledkami a rozpätiami výnosov poskytujú priestor pre identifikáciu dobrých investičných príležitostí.

3 ÚDAJE A VÝPOČTY

OECD pravidelne poskytuje historické a prognostické ročné makroekonomické údaje, ktoré opisujú:

- dopyt a produkciu,
- mzdy, náklady, nezamestnanosť a infláciu,
- kľúčové údaje ponuky,
- úspory,
- fiškálnu bilanciu a údaje o verejnom dlhu
- úrokové miery a výmenné kurzy,
- zahraničný obchod a platobnú bilanciu,
- iné základné údaje,

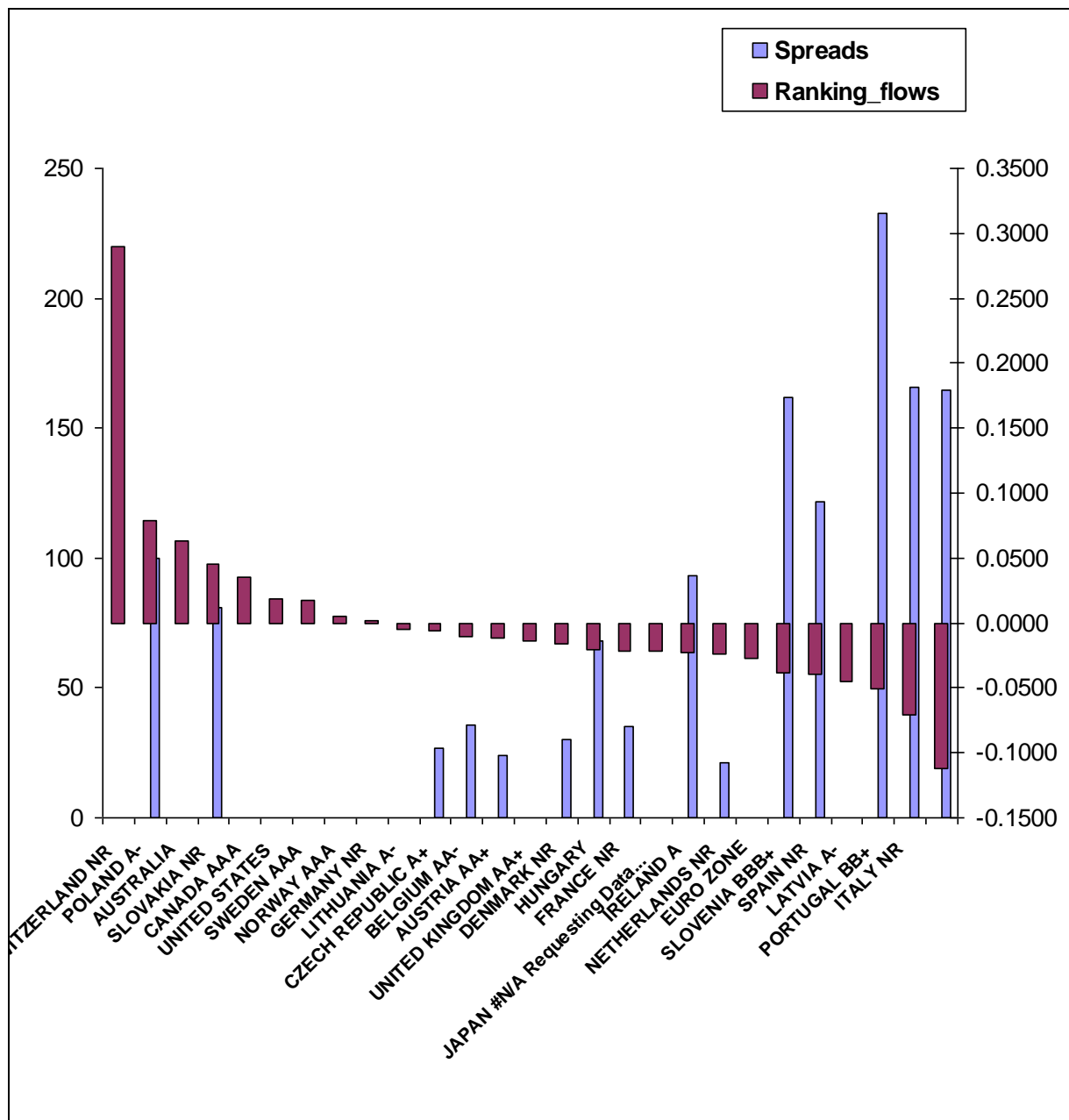
Tieto údaje pre obdobie od roku 2008 do roku 2015 boli použité pre konštrukciu nasledujúcich makroekonomických kritérií:

1. hrubí pridaná hodnota v stálych cenách,
2. súkromná konečná spotreba v stálych cenách,
3. hrubá tvorba fixného kapitálu v stálych cenách,
4. miera nezamestnanosti,
5. harmonizovaný index spotrebiteľských cien
6. miera úspor ako percento hrubého disponibilného dôchodku,
7. národný dôchodok v bežných trhových cenách,
8. vládna finálna spotreba v stálych cenách,
9. rozpočtový deficit ako % HDP,
10. implicitná úroková miera,
11. všeobecné vládne výdavky ako % HDP
12. nominálna dlhodobá úroková miera,
13. celkový export tovarov a služieb,
14. celkový dovoz tovarov a služieb,
15. bilancia bežného účtu ako % HDP,
16. rast trhu vývozu tovarov a služieb.

Systém na podporu rozhodovania vyvinutý v excelovskom prostredí realizuje metódu Promethee II vrátane analýzy senzitivnosti pre váhy kritérií. Umožňuje špecifikovať:

- obdobie pre kritériá úrovne (v ilustrácii 2008 – 2015)
- obdobie pre kritériá trendu (2012 – 2015),
- krajiny zo zadaného zoznamu,
- kritériá zo zadaného zoznamu spolu s priradením váh,
- riadiace parametre pre metódu Promethee II.

Ako ilustráciu uvádzame výsledky na obr.1, v ktorom sa konfrontujú hodnoty čistých tokov jednotlivých krajín s výnosovými rozpätiami.



Obrázok 1: Výsledky pre rozpätia vzhľadom k nemeckým dlhopisom

Použitá literatúra

1. MLYNAROVÍČ, V. 1998. Modely a metódy viackriteriálneho rozhodovania. Ekonóm, Bratislava, 1998

Kontaktné údaje

doc. Ing. Vladimír Mlynarovič, CSc.

Fakulta sociálnych a ekonomických vied Univerzity Komenského v Bratislave

Mlynské luhy 4, 821 05 Bratislava

email: vladimir.mlynarovic@fses.uniba.sk

NOVÝ SOFTWARE PRE METÓDY TRIEDY ELECTRE

NEW SOFTWARE FOR ELECTRE METHODS

Michal Pieter

Abstrakt

Metódy triedy ELECTRE predstavujú populárny nástroj riešenia niektorých úloh v teórii rozhodovania. Tento príspevok predstavuje v stručnosti nový software, nazvaný ELIA, ktorý bol vyvinutý ako súčasť autorovej diplomovej práce. Tento software umožňuje riešenie úloh pomocou metódy ELECTRE IS poskytujúcej určenie množiny kompromisných variantov a metód ELECTRE II, III a IV, poskytujúcich zoradenie variantov od najvhodnejšieho po najmenej vhodný. Súčasťou textu je taktiež stručný popis základných princípov, na ktorých sú implementované metódy založené, pričom záujemcovia o bližšie preskúmanie sú odkazovaní na diplomovú prácu a primárnu literatúru.

Kľúčové slová: teória rozhodovania, VHV, ELECTRE, software, ELIA

Abstract

ELECTRE methods are a popular tool for solving certain decision-making problems. This paper presents new software called ELIA that was developed as part of the authors master's thesis. It supports the ELECTRE IS method, selecting a set of compromise actions, as well as ELECTRE II, III and IV methods for ordering the actions from best to worst. A brief summary and working principles for all of these methods is provided, including references to the master's thesis and primary sources for more detailed study.

Keywords: decision theory, MCDM, ELECTRE, software, ELIA

1 ÚVOD

V rámci teórie rozhodovania je možné identifikovať viacero hlavných prístupov. Oblasť viackriteriálneho rozhodovania (multi-criteria decision making, MCDM) sa zaoberá problémami, kedy je vhodné analyzovať množinu možných rozhodnutí, väčšinou výberom toho najvhodnejšieho, na základe daných kvantitatívnych alebo kvalitatívnych kritérií. V prípade, že je táto množina diskretná, jednotlivé rozhodnutia sú nazývané variantami a jedná sa o úlohu viackriteriálneho hodnotenia variantov (VHV, multi-criteria evaluation). Medzi tieto patrí aj rodina úloh ELECTRE, vyvinutá koncom šesťdesiatych rokov Bernardom Royom a v nasledujúcich dekádach postupne rozširovaná najmä ním a jeho spolupracovníkmi.

Práve prehľad metód triedy ELECTRE, ich modifikácií a rozšírení, ako aj oblastí ich aplikácie a súvisiacich programových prostriedkov bol témou diplomovej práce Metody triedy ELECTRE: Teorie a aplikace [1], ktorá bola odovzdaná v júni 2016 a obhájena v septembri toho istého roku autorom príspevku. Jedným z cieľov danej práce bolo taktiež vytvorenie úplne nového software, ktorý by umožňoval riešenie úloh VHV pomocou niekoľkých zo základných metód triedy ELECTRE. Stručné predstavenie práve týchto metód, ako aj samotného novovytvoreného software – nazvaného ELIA – je pritom zámerom tohto príspevku. Vzhľadom na rozsah problematiky a obmedzený priestor sú viaceré detaily

vynechané, záujemcov o komplexnejší pohľad je preto možné odkázať na diplomovú prácu a na primárnu literatúru.

2 METÓDY TRIEDY ELECTRE

Metódy triedy ELECTRE patria do skupiny metód, ktoré sú výpočtovo založené na preferenčných reláciách medzi jednotlivými variantmi. Každý variant má podľa každého kritéria priradenú určitú číselnú (kriteriálnu) hodnotu, ktorá vyjadruje vhodnosť daného variantu, a to buď v kardinálnom alebo ordinálnom zmysle. Párovým porovnaním variantov je potom možné určiť jeden z troch tradičných preferenčných vzťahov – pre varianty a a b sú to:

- aPb : variant a je (striktne) preferovaný pred variantom b ,
- bPa : variant b je (striktne) preferovaný pred variantom a ,
- aIb : varianty a a b sú navzájom indiferentné.

Metódy triedy ELECTRE však tieto preferenčné vzťahy rozširujú zavedením viacerých prahových hodnôt. Ich zmyslom je v modeli lepšie zachytiť nepresnú povahu reálneho systému a neistotu panujúcu pri stanovení spomenutých preferenčných vzťahov. Kľúčovým konceptom je pritom vzťah tzv. nadradenosti (outranking) dvoch variantov, zapisovaný ako aSb , ktorý vyjadruje, že variant a je možné považovať za aspoň taký dobrý, ako variant b . Obmedzený vzťah nadradenosti aS_jb potom hovorí to isté, ale len pri porovnaní podľa kritéria j . Keď je kriteriálna hodnota variantu a podľa kritéria j značená ako $g_j(a)$, je možné s využitím už zmienovaných prahových hodnôt definovať preferenčné relácie aP_jb , aI_jb a aS_jb ako

$$\begin{aligned}aI_jb &\Leftrightarrow g_j(a) - g_j(b) \leq q_j, \\aS_jb &\Leftrightarrow g_j(a) \geq g_j(b) - q_j, \\aP_jb &\Leftrightarrow g_j(a) > g_j(b) + p_j.\end{aligned}$$

Hodnota q_j sa nazýva prah indiferencie (indifference threshold) a vyjadruje maximálny rozdiel kriteriálnych hodnôt variantov tak, aby ešte mohli byť podľa kritéria j považované za indiferentné. Obdobne, hodnota p_j sa nazýva prah preferencie (preference threshold) a predstavuje minimálny rozdiel, kedy je ešte jeden variant považovaný za preferovaný pred druhým. Tieto hodnoty môžu byť konštantné, prípadne môžu byť funkčnou hodnotou kriteriálnej hodnoty variantu, kedy sa označujú $q_j(g_j(a))$, resp. $p_j(g_j(a))$. [2, s. 51-52, 57-58]

2.1 ELECTRE IS

Metóda ELECTRE IS (jedna-es) je rozšírením metód ELECTRE I a Iv (jedna-vé), pričom všetky si kladú za cieľ z množiny variantov určiť jeden či viacej kompromisných, ktoré je možné považovať za najvhodnejšie. Jej princíp spočíva v agregácii uvedených čiastkových vzťahov do obecného vzťahu nadradenosti aSb pre každú dvojicu variantov. Pre každú takúto dvojicu sa snaží roztriediť množinu kritérií do troch disjunktných podmnožín – koalícií – podľa toho, či dané kriteriálne hodnoty, v kombinácii so stanovenými prahmi preferencie a indiferencie, podporujú – pre dané kritérium j platí aS_jb – alebo odporujú – platí bP_ja – tvrdeniu, že platí vzťah aSb , prípadne sa k jeho platnosti nedokážu jasne vyjadriť. [2, s. 57-58]

Väčšina metód triedy ELECTRE využíva kardinálnu informáciu o relatívnej dôležitosti jednotlivých kritérií, teda tieto majú pridelené určité váhy. ELECTRE IS je jednou z nich a silu koalícií tak obecné určuje ako súčet váh kritérií, ktoré patria do danej koalície. V metóde ELECTRE IS vyjadruje súčet (pri ignorovaní korekčného koeficientu [1, s. 16]) síl oboch koalícií, ktoré nie sú nesúhlasné s tvrdením aSb tzv. index súhlasu. V prípade, že tento prekračuje určitú pevnú hodnotu – úroveň súhlasu – zadanú rozhodovateľom, je možné považovať tvrdenie aSb za platné. [2, s. 60-61] Umožňuje však, rovnako ako metóda ELECTRE Iv, ktorú rozširuje, aby dostatočne veľký rozdiel hodnôt pre nejaké kritérium

predstavoval akési veto pre platnosť tohto tvrdenia. Daný minimálny rozdiel sa preto nazýva prah veta (veto threshold) a rozhodovateľ ho môže stanoviť pre jednotlivé kritéria individuálne buď vo forme konštanty alebo funkcie kritériálnej hodnoty. [3, s. 142]

Keď je stanovená platnosť vzťahu aSb pre každú dvojicu variantov (vzťah aSa platí vždy), je možné odvodiť platnosť vzťahov preferencie, indiferencie a neporovnateľnosti podľa [2, s. 54]

$$\begin{aligned} aSb \wedge \neg bSa &\Leftrightarrow aPb, & \forall a, b \in A, & a \neq b, \\ aSb \wedge bSa &\Leftrightarrow aIb, & \forall a, b \in A, & a \neq b, \\ \neg aSb \wedge \neg bSa &\Leftrightarrow aNb, & \forall a, b \in A, & a \neq b. \end{aligned}$$

Tieto vzťahy sú následne zachytené vo forme orientovaného grafu. Pri získaní konečnej výstupnej informácie sa využije koncept tzv. jadra grafu [4, s. 183], ktoré predstavuje množinu kompromisných variantov. Procesom komunikácie s rozhodovateľom a postupnou zmenou prahov a úrovní je potom možné získať práve jeden kompromisný variant, ktorý je tak možné považovať za najvhodnejší. [3, s. 140-41]

2.2 ELECTRE II a III

Metódy ELECTRE II a III sa odlišujú od predchádzajúcej najmä tým, že poskytujú ako výstupnú informáciu zoradenie variantov od najvhodnejšieho po najmenej vhodné. Metódy ELECTRE II je rozšírením metódy ELECTRE IS, kedy rozdeľuje vzťah nadradenosti aSb na dva, s ktorými sú potom spojené aj dve úrovne súhlasu. Tieto vzťahy majú prívlastok silný – aS^1b – a slabý – aS^2b – pričom klasický vzťah nadradenosti je potom predefinovaný ako

$$aSb \Leftrightarrow aS^1b \wedge aS^2b.$$

Odhalením platnosti vzťahov silnej a slabej nadradenosti pre každú dvojicu variantov je možné zostrojiť dva orientované grafy – jeden, kde hrany predstavujú vzťah výhradne silnej nadradenosti a druhý, kde orientovaná hrana vyjadruje platnosť aspoň jedného z nich. Oba grafy sú následne očistené od prípadných cyklov a pomocou iteratívneho procesu je získaná dvojica čiastkových zoradení. Konečné poradie variantu je potom získané ako prienik alebo alternatívne ako aritmetický priemer jeho oboch čiastkových poradí. [3, s. 144-45]

Metóda ELECTRE III predstavuje modifikáciu metódy ELECTRE II, ale taktiež ELECTRE IS, kde modifikuje index súhlasu na tzv. index dôveryhodnosti, ktorý je však závislý aj na sile nesúhlasnej koalície s tvrdením aSb . Taktiež umožňuje modelovanie čiastočného efektu veta. Získaný index dôveryhodnosti vyjadruje akúsi mieru platnosti tvrdenia aSb , čo je v kontraste s predchádzajúcimi metódami, ktoré operovali len s dichotomickými stavmi platí–neplatí. [3, s. 145] Z týchto indexov pre každú dvojicu variantov získava ich čiastkové zoradenie pomocou procesu tzv. destilácie, ktorá vyžaduje stanovenie dvoch parametrov rozhodovateľom. Tie je možné interpretovať ako jemnosť rozlíšenia variantov v konečnom poradí. To sa získa obdobne ako pri metóde ELECTRE II kombináciou poradí čiastkových. [5, s. 3]

2.3 ELECTRE IV

Metóda ELECTRE IV je netradičná, pretože nepracuje s váhami kritérií a nevyžaduje tak od rozhodovateľa ich stanovenie. Vychádza z metódy ELECTRE III a podobne ako metóda ELECTRE II tiež rozdeľuje vzťah nadradenosti, tentoraz však hneď na 5 čiastkových vzťahov aS_1b až aS_5b . Tie sú pritom hierarchicky zoradené a platnosť jedného implikuje zároveň platnosť voči nemu silnejších, teda viacej obmedzujúcich. Pre každú dvojicu variantov zatrieduje množinu kategórií do jednej zo siedmich disjunktných skupín, pomocou čoho následne skúma platnosť jednotlivých vzťahov. [1, s. 22-23] [6, s. 217] Každému z nich pritom zodpovedá určitá klesajúca hodnota indexu dôveryhodnosti voči tvrdeniu aSb . Takto

získané hodnoty indexov sú následne pretavené procesom destilácie do končného zoradenia variantov, rovnako ako v prípade metódy ELECTRE III. [3, s. 147]

3 SOFTWARE ELIA

Programových prostriedkov, ktoré by podporovali niektorú z metód triedy ELECTRE, je v súčasnosti k dispozícii na trhu relatívne málo. Väčšina z nich pochádza práve z laboratória LAMSADE, prípadne bola vytvorená v spolupráci s ním. Menovite sa jedná o programy ELECTRE IS, ELECTRE III-IV a ELECTRE Tri. Tieto boli vyvinuté v 90-tych rokoch a z toho dôvodu trpia hneď niekoľkými nedostatkami. Používateľské rozhrania sú značne zastarané, čo môže sťažovať prácu používateľom, ktorí sú zvyknutí na moderné aplikácie, ktoré väčšinou dodržiavajú dnes už obecné normy používateľského dizajnu. Napriek tomu, že tí zasvätení do metód ELECTRE nemajú po určitej dobe väčší problém sa v nich orientovať, z hľadiska praktického použitia už je situácia iná. Bez znalosti terminológie a teórie jednotlivých metód je ťažké predstaviť si ich použitie. Najväčší problém však predstavuje fakt, že žiaden z týchto programov nie je možné spustiť na moderných 64-bitových operačných systémoch, ktoré tvoria väčšinu na trhu a sú takmer výhradne inštalované na nových počítačoch. [1, s. 51-54]

Existujú aj ďalšie programy, napríklad IRIS [7], SANNA [8] a niektoré ďalšie, tie však nepodporujú niektoré zo základných metód predstavených v kapitole 2. Nový software, nazvaný ELIA, ponúka riešenie spomenutých problémov. Je ho možné bez problémov používať na moderných operačných systémoch Microsoft Windows, ponúka prehľadné a intuitívne ovládanie a v súčasnej verzii podporuje štyri základné modely triedy ELECTRE. Zároveň bol koncipovaný tak, aby jeho funkcionality bolo možné čo najjednoduchšie ďalej rozširovať.

Ako už bolo spomenuté v úvode, program ELIA bol vytvorený v rámci diplomovej práce, ktorá bola obhájená v septembri roku 2016. Od tej doby boli opravené nahlásené chyby a zapracované pripomienky k programu ELIA z posudkov práce. Napriek tomu, že program je hotový a funkčný, niektoré aspekty je ešte vhodné odladiť, ako aj odstrániť chyby odhaliteľné až pri jeho širšom použití. Následne bude program zverejnený na stránkach Katedry ekonometrie VŠE (ekonometrie.vse.cz), kde bude voľne prístupný na stiahnutie pre nekomerčné využitie. Termín sprístupnenia sa pritom očakáva v priebehu prvých mesiacov roku 2017.

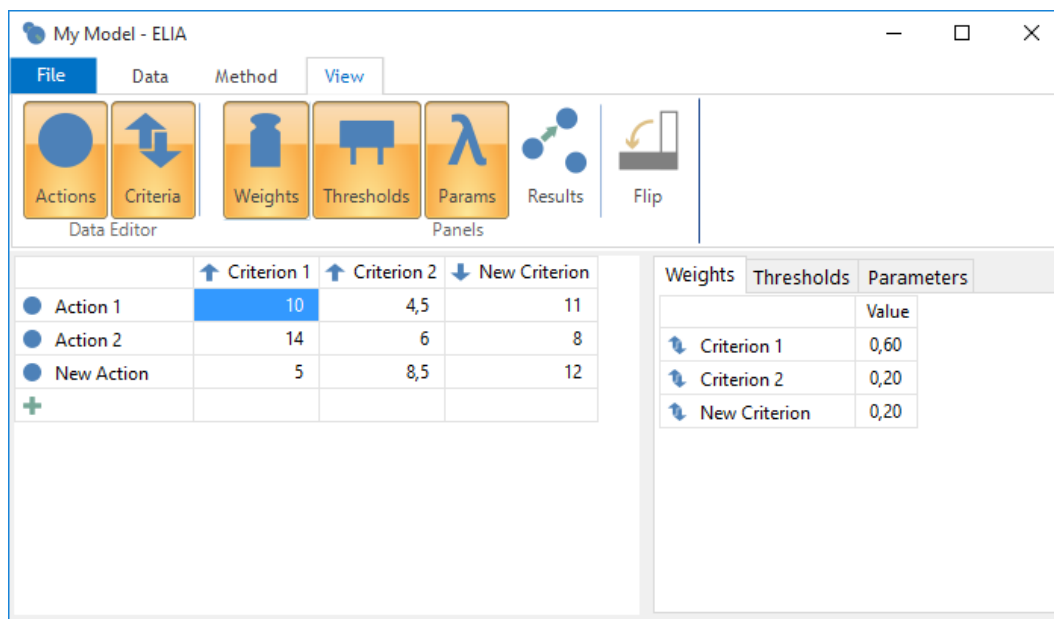
V nasledujúcich dvoch kapitolách je možné nájsť stručný prehľad funkcionality, niektorých obmedzení a možností ďalšieho vývoja a rozvoja aplikácie ELIA. Detailnejšiu analýzu, proces vývoja a internú architektúru, ako aj implementáciu jednotlivých algoritmov vrátane pseudokódu je potom možné nájsť v samotnej diplomovej práci [1, s. 57-72].

3.1 Prehľad funkcionality

Program ELIA je možné spustiť na operačných systémoch Microsoft Windows, a to vo verzii Vista a novšie. Na jeho spustenie je nutné mať rozhranie Microsoft .NET Framework vo verzii aspoň 4.5.1, ktoré je voľne dostupné z oficiálnych stránok Microsoft-u a program pri spustení upozorní, ak je ho nutné nainštalovať. Na operačnom systéme Windows 8.1 a novších je už priamo jeho súčasťou a nie je od používateľa vyžadovaná žiadna ďalšia operácia. Program samotný nie je vôbec nutné inštalovať, stačí ho len spustiť otvorením súboru ELIA.exe.

Práca s programom prebieha v jednom okne, ktoré je možné vidieť napríklad na obr. 1. Toto je rozdelené niekoľko hlavných častí. Tabuľka s dátami v dolnej časti má centrálnu pozíciu

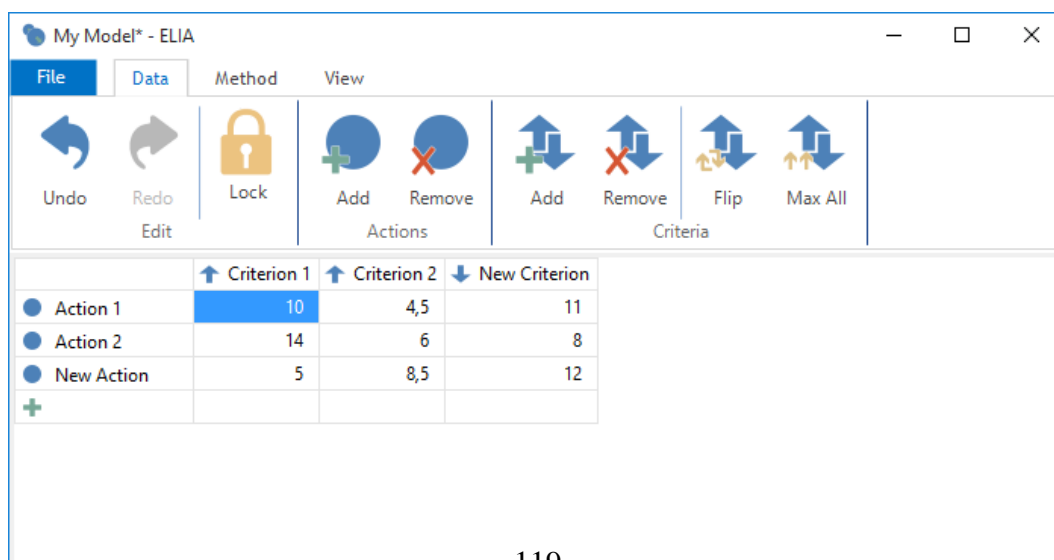
a umožňuje zobrazit' aj rozsiahly súbor dát. Lišta nástrojov sa nachádza v hornej časti a obsahuje tri záložky – Data, Method a View (program ELIA je v súčasnej verzii len



Obr. 1 – Okno aplikácie ELIA, záložka View

v anglickom jazyku). Lištu nástrojov je možné schovať opätovným kliknutím na práve zvolenú záložku. V dolnej časti je taktiež umiestnený panel, ktorý slúži na zadávanie parametrov modelu a zobrazenie výsledkov. V lište nástrojov, na záložke View je možné zvoliť, ktoré záložky sa majú zobrazovať v paneli. Ak je nutné rozšíriť pracovný priestor tabuľky dát, je možné panel kedykoľvek schovať (klávesová skratka Ctrl+T) a znova zobrazit' v pôvodnej podobe. Je ho taktiež možné presunúť do spodnej časti okna pomocou tlačidla Flip (tiež Ctrl+Shift+T). Ako je možné vidieť z obr. 1, tlačidlá sú veľké, prehľadne usporiadané a jasne popísané. Pri podržaní myši nad tlačidlom sa taktiež zobrazí text s krátkym popisom a prípadne tiež nápovedou.

Pri otvorení programu sa naskytne používateľovi pohľad na okno s už otvoreným prázdny príkladom, tak ako je vidno na obr. 2. Ten je možné prispôbiť vlastným potrebám, prípadne je možné vytvoriť nový prázdny model. To môže používateľ cez hlavné menu, kliknutím na tlačidlo File v lište nástrojov a zvolením možnosti New (Ctrl+N). Tam je v súčasnej verzii už možné taktiež zvoliť rozmer novovytvoreného modelu, teda počet variantov a kritérií. Možné je taktiež otvoriť už existujúci model cez tlačidlo Open (Ctrl+O), pričom dátové súbory majú príponu .elm. Pod rovnakým typom je samozrejme možné aj uložiť existujúci súbor tlačidlom Save (Ctrl+S), prípadne Save As (Ctrl+Shift+S).



Obr. 2 – Tabuľka s kritériálnymi hodnotami úlohy, záložka Data

Práca s dátami je intuitívna a jednoduchá, údaje sa zadávajú rovnako ako v bežných tabuľkových kalkulátoroch. Varianty (action) sú reprezentované riadkami a kritériá (criterion) stĺpcami tabuľky. Ak chce používateľ zmeniť ich názvy, musí vykonať dvojklik nad danou hlavičkou a zadať nový názov. Jednorazové kliknutie totiž riadok, stĺpec, prípadne viacero z nich zvýrazní. Následne je ho, resp. ich možné cez tlačidlo Remove v záložke Data na lište nástrojov odstrániť, alebo v prípade kritérií zmeniť ich orientáciu z maximalizačného na minimalizačné a naopak pomocou tlačidla Flip. V nastaveniach (Options) v hlavnom menu je možné zvoliť či a ako sa majú pri zmene orientácie upraviť údaje v relevantných stĺpcoch. Pridávanie variantov a kritérií je taktiež jednoduché. Stačí použiť zodpovedajúce tlačidlo Add v lište nástrojov, alebo v prípade variantov jednoducho vyplniť hodnoty v poslednom riadku. Nakoniec, dáta je možné vložiť z tabuľkového kalkulátoru, prípadne z textového súboru (tzv. tab-separated values) aj priamo. To je možné buď pretiahnutím označených dát z externého zdroja do oblasti tabuľky v programe ELIA, alebo kopírovaním, označením celej tabuľky kliknutím na ľavú hornú bunku a stlačením klávesovej skratky Ctrl+V. V oboch prípadoch je nutné dodržať formát riadky–varianty, stĺpce–kritériá a celú operáciu potvrdiť.

Po vyplnení tabuľky kritériálnych hodnôt je pred samotným výpočtom nutné zvoliť konkrétnu metódu a vyplniť zodpovedajúce parametre. Na záložke Method zvolí používateľ jednu z metód – v súčasnej verzii sú podporované metódy ELECTRE IS, II, III a IV. Podľa konkrétnej metódy je nutné vyplniť hodnoty parametrov vo výsuvnom paneli. Napríklad pre metódu ELECTRE IS sa jedná o váhy kritérií v paneli Weights, úroveň súhlasu v paneli Parameters a jednotlivé prahy v paneli Thresholds. Tie sa zadávajú vo forme lineárnej funkcie $f(x)=\alpha+\beta x$ prostredníctvom funkčných parametrov α (y-intercept) a β (gradient). Pred spustením výpočtu sa automaticky skontroluje správnosť a konzistentnosť zadaných údajov, tú je možné tiež kedykoľvek skontrolovať cez tlačidlo Verify na záložke Method. Prípadné problémy sú takto vymenované.

Výpočet samotný sa spúšťa tlačidlom Calculate na záložke Method. Keď úspešne prebehne, otvorí sa automaticky panel s výsledkami, a to podľa zvolenej metódy – kompromisné varianty pre metódu ELECTRE IS a zoradenie variantov pre metódy ELECTRE II, III a IV. Ako je možné vidieť aj na obr. 3, výsledky sú farebne kódované, v prípade zoradenia gradientom od zelenej pre najvhodnejší varianty až po červenú pre ten najmenej. Viacero variantov pritom môže zdieľať poradové umiestenie a odtieň. Takéto odlišenie je možné

	↑ Criterion 1	↑ Criterion 2	↓ New Criterion
● Action 1	10	4,5	11
● Action 2	14	6	8
● New Action	5	8,5	12
+			

Results

Sorting into categories

Efficient actions:
Action 2

Inefficient actions:
Action 1
New Action

Obr. 3 – Výsledok výpočtu metódou ELECTRE IS

vypnúť v nastaveniach.

3.2 Obmedzenia a ďalší vývoj

Vývoj aplikácie ELIA nebol hlavným ani jediným cieľom diplomovej práce, aj z toho dôvodu museli byť pri jej vývoji učené viaceré kompromisy. Aspoň v jej súčasnej verzii je teda možné identifikovať niekoľko obmedzení. Tie však zároveň predstavujú príležitosť pre jej ďalšie zdokonaľovanie a rozširovanie funkcionality. Ako všetok nový software, aj program ELIA nutne obsahuje chyby či nedostatky, ktoré sa v priebehu vývoja nepodarilo odhaliť. Tieto budú po nahlásení postupne sa zväčšujúcou skupinou používateľov priebežne odstraňované a takto zdokonalené verzie postupne uverejňované v priebehu ďalších mesiacov a rokov.

V súčasnosti podporuje štyri zo základných metód triedy ELECTRE, nepochybne by preto bolo vhodné pridať aj ďalšie, medzi ktorými by určite nesmela chýbať metóda ELECTRE TRI, prípadne jej modifikácie. Celý program bol už od začiatku vývoja koncipovaný tak, aby dorábanie nových metód bolo čo najjednoduchšie a zároveň aby bola zaručená kompatibilita s modelmi vytvorenými predošlými verziami. Výskum v oblasti metód ELECTRE pritom poskytuje stále sa zväčšujúce množstvo implementovateľných metód a ich rozšírení.

Momentálne je používateľské rozhranie riešené spôsobom, ktorý vyžaduje aspoň určitú znalosť jednotlivých metód triedy ELECTRE. Zároveň výstup výpočtu je v súčasnosti veľmi jednoduchý. Zámerom autora je však, aby mohol byť program využívaný ako pre riešenie praktických rozhodovacích úloh zodpovednými osobami, taktiež však ako učebná pomôcka v rámci teórie rozhodovania. Kľúčom k dosiahnutiu oboch cieľov je ďalšie zjednodušenie rozhrania a zároveň rozšírenie poskytovaných výstupných informácií. Tie môžu mať formu správy, rôznych tabulovaných údajov, grafickej reprezentácie výsledkov, prípadne orientovaných grafov. Pre účely učebnej pomôcky je možné si predstaviť možnosť krokovania výpočtu a previazanie výstupnej informácie s interným algoritmom v reálnom čase.

Použitá literatúra

1. PIETER, M. *Metody třídy ELECTRE: Teorie a aplikace*. Praha: 2016. Diplomová práca. Vysoká škola ekonomická v Praze.
2. ROY, B. The outranking approach and the foundations of electre methods. 1991, **31** (1), 49-73. ISSN: 0040-5833, 1573-7187.
3. JOSÉ FIGUEIRA, V. M. Electre Methods. 2004 roč. LXXVIII, s. 133-53.
4. SCHMIDT, G. a T. STRÖHLEIN. *Relations and Graphs: Discrete Mathematics for Computer Scientists*. New York: Springer Science & Business Media, 2012. ISBN: 978-3-642-77968-8.
5. REDONDO PENA, R. et al. *Use and Evaluation of Electre III/IV*. Tarragona: Universitat Rovira i Virgili, 2007.
6. HOKKANEN, J. a P. SALMINEN. ELECTRE III and IV Decision Aids in an Environmental Problem. 1997, **6** (4), 215-26. ISSN: 1099-1360.
7. DIAS, L. C. GOMES DA SILVA a V. MOUSSEAU. *IRIS - Interactive Robustness analysis and parameters' Inference for multicriteria Sorting problems ...* [Dokument]. 2002. ISSN 1645-4847.
8. JABLONSKÝ, J. a P. URBAN. *MS Excel based system for multicriteria evaluation of alternatives*. Praha: 1998. University of Economics Prague, Department of Econometrics.

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: xpiem02@vse.cz

DEFINICE ZÁKLADNÍHO GRAVITY MODELU A JEHO KOMBINACE S MODELY PROSTOROVÉ AUTOKORELACE

DEFINITION OF BASIC GRAVITY MODEL AND ITS COMBINATION WITH SPATIAL MODELS

Tatiana Polonyankina

Abstrakt

Príspevek se venuje gravity modelu, ktorý je zaujímavý svojí analogií s Newtonovým gravitačným zákonom. Logika gravitace se v modelu přeneseně používá pro popis prostorové interakce mezi ekonomickými jednotkami. Cílem je definovat kombinaci základního gravity modelu s modely prostorové autokorelace. První sekce představuje gravity model, stručně sumarizuje historii a ukazuje možnosti použití modelu v různých ekonomických odvětvích. Další část definuje gravity model a ukazuje jeho kombinaci s modely prostorové autokorelace. Poslední část je simulací aplikace modelu kombinujícího gravity model a model prostorové autokorelace na reálná data.

Klíčová slova: prostorové modely, gravity model, prostorová autokorelace

Abstract

This work is introducing the gravity model which derives its name and basic idea from Newton's gravity law. The idea of gravitation is used to describe spatial interaction between economic cells. The main aim is the investigation of the link between the spatial modelling and the basic gravity model. Section motivation is introducing the gravity model, explaining a bit of history and also summarising the fields of applying the model. The next session is setting the theoretical definition of the gravity model and showing the connection with the spatial models. The last part consists of simulation of migration analysis with a model which combines the gravity model and a spatial model.

Keywords: spatial models, gravity model, spatial autocorrelation

This work was supported by IGA F4/73/2016, University of Economic, Prague.

1 MOTIVATION

Spatial interaction is a term used to describe any movement over space that is caused or forced by human interaction or any other human activity. The common example can be any kind of migration (from workers going every day to a city for work to immigrants living in a host county) but it can also be a commodity flow, international trade or knowledge exchange.

The gravity model is used to explain this spatial interaction between countries since it is taken into account a geography aspects. The model derives its name and a basic idea from physics. There is an analogy to the gravitation interaction between planetary bodies where two main influences are the distance between the bodies and their scale impact. The scale can be measured in absolute values (size) or in a relative values depending on a field of model application (e.g. GDP as a measure of economic scale). Distance can be also defined as an absolute location – distance between two places in kilometres/miles; or as a relative location –

where locations with the same absolute distance can have different access to a shopping centre, education, museums, access to a job opportunities, information even access to water.

The use of gravity models for spatial interaction analysis is summarised in Constantin (2004), where he talks about the market area boundary analysis, gravity models for commodity flows analysis and gravity models for studying the migration flows. The classical gravity model was firstly used to explain trade flows by Tinbergen (1962). Later on the model was widely used for international trade analysis (Anderson (1979)).

The other macroeconomic field where the gravity model is widely used is analysis of foreign direct investment (FDI). Egger (2004) examines the effects of the EU integration in the 90's on the FDI relations in Europe using the gravity model. He found that the completion of the Single Market, the 1995 enlargement and the next enlargement are characterized by a substantial, positive anticipation effect in the period between the announcement and the formal establishment of each integration step. Braconier (2005) explores how wage costs for high skilled and less skilled labour in host countries are affected by FDI using the gravity model.

For migration flows was the gravity model used firstly by Ravenstein (1889) for UK. Metulini (2013) is estimating a Spatial Gravity Trade for panel of the OECD countries showing that Gravity model brings better results than classical ordinary least squares regression (OLS) where estimations are not unbiased and suffer from endogeneity problem.

2 SPATIAL ECONOMETRICS THEORY AND GRAVITY MODEL DEFINITION

The basic gravity model says that interaction T_{ij} between two regions i and j is a function of P_i the scale impact of region i , P_j the scale impact of region j and d_{ij} distance between two regions and the gravity constant G .

$$T_{ij} = G \frac{P_i^a P_j^b}{d_{ij}^c} \quad (1)$$

To estimate this equation we need to log-linearize the function (1) and add an error term. To do so we receive:

$$\log T_{ij} = \log G + a \log P_i + b \log P_j - c \log d_{ij} + \varepsilon_{ij} \quad (2)$$

For easier work with the equation we do a substitution in the equation (2) and receive:

$$y_{ij} = \alpha + \beta_d x_{di} + \beta_o x_{oi} + \gamma d_{ij} + \varepsilon_{ij} \quad (3)$$

We continue with a model expansion when we add more regions. The model now has total number of n regions. The interaction between regions is described with n by n square matrix of interregional flow each of n origin regions to each of the origin destination regions. This matrix we can restructure on n^2 by 1 vector y by converting columns of the matrix into a variable vector. A typical least-square gravity model will be:

$$y = \alpha I + X_d \beta_d + X_o \beta_o + D\gamma + \varepsilon, \quad (4)$$

Where:

- I is $n^2 \times 1$ intercept vector,
- X_d is $n^2 \times k$ vector of destinations region scale characteristics,
- X_o is $n^2 \times k$ vector of origin regions scale characteristics,

- D is $n^2 \times 1$ vector of origin – destination regions distances,
- $\varepsilon \sim N(0, \sigma^2 I_n)$,
- $\alpha, \beta_d, \beta_o, \gamma$ are parameters to be estimated.

The traditional gravity model can be combined with spatial models, where spatial matrix W is used to describe geographic dependencies. The spatial matrix W can be created taking into account just influence between neighbours or looking how regions are related thanks to distance between each other.

$$W = (cell_{i,j}), \quad i = 1, \dots, n, j = 1, \dots, n \quad (5)$$

with 0 or 1 by contiguous between cells, or with distance between cells.

Three types of basic spatial regression models can be used: spatial lag regression (SAR), spatial error regression (SEM) and spatial autocorrelation model (SAC). For more see LeSage (2009). The most complex is combination of SAR and SEM, called Spatial Autocorrelation Model (SAC) and it contains spatial dependence in both the dependent variable and the disturbances. The model is a logic decision when previous models (SAR and SEM) have both spatial parameters significant. The definition of the model is follows:

$$y = \rho W y + X \beta + u, \text{ where } u = \sigma W u + v, \quad (6)$$

$$(1 - \rho W) y = X \beta + \frac{v}{1 - \sigma W}, \quad (7)$$

$$v \sim N(0, \sigma^2 I_n). \quad (8)$$

What is the purpose of a combination of models? The gravity model takes into account origin and destination regions relations when spatial models capture geographic dependence between regions (one observation is one region). To combine the methods we need to construct an expanded spatial matrix when $n^2 \times 1$ vector reflects flows from all origins to all destinations. This can be done by using $I_n \times W$ to create $n^2 \times n^2$ row standardized spatial weight matrix:

$$W_o = \begin{pmatrix} W & 0 & \dots & 0 \\ 0 & W & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & W \end{pmatrix} \quad (9)$$

Using this matrix in spatial lag of the dependent variable $W_o y$ we can capture “origin-based” spatial relations. That means we can capture the force that leads to flow from any origin region to a destination region can create similar flows from neighbours of the origin region to the same destination region.

The second type of spatial dependence that can arise is “destination-based.” Spatial $n^2 \times n^2$ matrix capturing this is W_d and is created as $W \times I_n$. That means we can capture the force leading to flow from any origin region to a destination region can create similar flows to neighbours of the destination region.

The third dependence is combination of both effects described above =>”origin-destination-based” spatial relations. The expanded spatial matrix is defined as follows:

$$W_{o,d} = W_o \cdot W_d = (I_n \times W) \cdot (W \times I_n) = W \times W \quad (10)$$

Using matrices defined above we can create spatial regression models for origin-destination flows. Given a spatial weight matrix W_j the gravity regression model (4) can be extended to a SAR, SEM or SAC model specification as:

$$y = \rho W_j y + \alpha I + X_d \beta_d + X_o \beta_o + D\gamma + z, z = \sigma W_j u + \varepsilon, \quad (11)$$

W_j can be any of the spatial weight matrices W_o, W_d, W_{od} , where W_o is used for capturing “origin-based” spatial relations, W_d is used for capturing “destination-based” spatial relations and W_{od} for “origin-destination-based” spatial relations.

3 SIMULATION OF AN ANALYSIS

The simulation of a migration analysis is done in this section. Regarding to the field of application a scale for interaction measurement is a migration between countries, a distance is measured in absolute values and scale for size is measured as an economic attraction of a country - GDP per capita. The aim is to prepare data for a regression of the gravity model with “origin-based” spatial relations. The analysis of migration is done for three different countries in year 3000, where countries one and three aren't neighbours and country two is a neighbour of country one and country three. GDP per habitant for countries is: 26 000 dollars/habitant, 40 000 dollars/habitant and 42 000 dollars/habitant. Distance is measured as distance between centres of countries in thousands km and defined in the matrix D (Distance matrix)

Distance matrix			
	c1	c2	c3
c1	0	10	40
c2	10	0	30
c3	40	30	0

Distance 0 means the distance of one region to itself. Computational problems can occur with this definition of the distance matrix, where observations with 0 are evaluated as missing observations by some computational SW. Possible solution is to change 0 for a small number. In this example we change it for 1. Migration between an origin and a destination country is defined below (in thousands).

Migration matrix			
	c1	c2	c3
c1	400	150	50
c2	300	600	100
c3	250	150	450

In the first step we start with the basic gravity model (1), where T_{ij} is migration between two countries i and j, P_i is GDP of the county i, P_j is GDP of the county j and d_{ij} is distance between the countries. Continuing the same way as is described at the theory part we create the model (4) where:

$$y = \alpha I + X_d \beta_d + X_o \beta_o + D\gamma + \varepsilon,$$

I=	1	Y=log	400	Xd=log	26 000	Xo=log	26 000	D=log	1
	1		150		40 000		26 000		10
	1		50		42000		26 000		40
	1		300		26 000		40 000		10
	1		600		40 000		40 000		1
	1		100		42000		40 000		30
	1		250		26 000		42000		40
	1		150		40 000		42000		30
	1		450		42000		42000		1

Now the data are prepared for a regression of the basic gravity model. Based in the economic theory we expect that β_d and β_o will be positive and γ will be negative.

In case we would like to continue and use the combination of gravity and spatial models we need to define the matrix W. Here we can choose if we want to use distances or just neighbour impact. In this example we are creating the matrix W with 0 or 1 by contiguous between cells and in the next step we do row standardization.

$$w = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}$$

Since we would like to create the model with capturing “origin-based” spatial relations we need to create W_o in size 9 x 9 as

$$W_o = \begin{pmatrix} W & 0 & \dots & 0 \\ 0 & W & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & W \end{pmatrix}$$

W _o =	0	1	0	0	0	0	0	0	0
	1/2	0	1/2	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0
	0	0	0	0	1	0	0	0	0
	0	0	0	1/2	0	1/2	0	0	0
	0	0	0	0	1	0	0	0	0
	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	1/2	0	1/2
	0	0	0	0	0	0	0	1	0

In this stage we have everything prepared for using the model (11) where $W_j = W_o$. As we can see the computational part can become quite large since the matrix W_j is in size $n^2 \times n^2$. The model is regressed usually using an algorithm of maximum likelihood.

4 CONCLUSION

The basic gravity model is interesting for its analogy with physics and Newton’s gravitation law. The main aim of the article was the investigation of the link between the spatial modelling and the basic gravity model. The theoretical part defined the gravity model and possible spatial models that can be combined with it to capture spatial autocorrelation between cells. Simulated data for three countries were used to simulate the application of gravity model combined with spatial models. It was found that a distance matrix should be used without 0 values since these observations can be evaluated as missing. The simulation is

concentrated on creation of spatial matrix where its size increases rapidly with number of countries used. Computational difficulties can arise as the number of observations becomes large. The next research should concentrate on modifications of the model and estimation techniques that should be used to avoid these difficulties.

References

1. ANDERSON J. A. 1979. A Theoretical Foundation for the Gravity Model. In: The American economic Review, roč. 69, s. 106-116.
2. R. METULINI. 2013. Spatial gravity models for international trade: a panel analysis among OECD countries. In: European Regional Science Association conference paper, č. ersa13p522.
3. EGGER, P., & PFAFFERMAYR. M. 2004. Foreign direct investment and European integration in the 1990s. in: The World Economy, roč. 27, č.1, s. 99-110.
4. BRACONIER, H., NORBRACK, P. J., & URBAN, D. 2005. Multinational enterprises and wage costs: Vertical FDI revisited. In: Journal of International Economics, roč.67, č.2, s. 446-470.
5. CONSTANTIN, D. L. 2004. The use of Gravity Models for spatial interaction analysis. In: Economy Informatics, roč. 1, č. 4, s. 116-118.
6. KINGSLEY, E., H., FORTHERINGHAM A.S. 1984. Gravity and Spatial Interaction Models, Vyd.: SAGE Publications, ISBN 0803923260
7. LESAGE J., PACE R. K. 2009. Introduction to spatial econometrics, Vyd.: CRC press, FL 33487-2742.
8. RAVENSTEIN, E. G. 1889. The laws of migration. In: Journal of the royal statistical society, roč. 52, čís. 2, s. 241-305.
9. TINBERG, J. 1962. Shaping the world economy; suggestions for an international economic policy. Vyd.: Books (Jan Tinbergen).

Kontaktní 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

ANALÝZA EKONOMICKÉHO RASTU A ENVIRONMENTÁLNYCH UKAZOVATEĽOV POMOCOU KLASTROVEJ ANALÝZY¹

CLUSTER ANALYSIS OF ECONOMIC GROWTH AND ENVIRONMENTAL INDICATORS

Marian Reiff, Zlatica Ivaničová

Abstrakt

Cieľom príspevku je analyzovať identitu (ktorá dáva do súvisu ekonomický rast, spotrebu energie a emisie oxidu uhličitého) pomocou klastrovej analýzy. Na základe dát z Eurostatu za rok 2014, konkrétne hrubý domáci produkt na osobu, spotrebu energie na osobu a emisií skleníkových plynov na osobu a za pomoci Wardovej metódy, zhlukujeme európske krajiny do homogénnych skupín s cieľom demonštrovať vzťah medzi ekonomickým rastom a emisiami skleníkových plynov.

KLúčové slová: emisie CO₂, klastrová analýza, Wardova metóda

Abstract

Aim of this paper is to analyze identity (equation relates economic growth, energy consumption and carbon dioxide emissions) using cluster analysis. Based on 2014 data from Eurostat, in particular, gross domestic product per capita, available energy for consumption per capita and greenhouse gas emissions per capita and utilizing Ward's method, a cluster of European countries - homogeneous groups were identified to demonstrate the relationship between economic growth and greenhouse gas emissions.

Keywords: CO₂ emissions, Cluster analysis, Ward's method

1 ÚVOD

Značný hospodársky rozvoj v priebehu dvadsiateho storočia viedol k zvýšeniu koncentrácie skleníkových plynov v atmosfére, predovšetkým oxidu uhličitého (CO₂). Celkové emisie oxidu uhličitého z fosílnych palív majú rastúci trend od roku 1900. Kaya (1990, 1997) sa zaoberal vplyvom ľudskej činnosti na zemskú atmosféru. Navrhol index, známy ako „Kaya identity“ ktorý dáva do súvisu celkové emisie CO₂ ktoré možno odhadnúť na základe štyroch faktorov: obsah uhlíka v spotrebovanej energii, energetická náročnosť hospodárstva, HDP na obyvateľa a veľkosť populácie. V príspevku analyzujeme tento vzťah pomocou klastrovej analýzy pomocou premenných: emisie CO₂ v tonách na obyvateľa, HDP na obyvateľa, vyprodukovaná energia na obyvateľa a podiel energie z obnoviteľných zdrojov na hrubej konečnej spotrebe energie. Cieľom nášho výskumu je identifikovať skupiny európskych krajín ktoré sú si navzájom podobné a to na základe vyššie uvedených ukazovateľov. Analýza vychádza z údajov Eurostatu za rok 2014. Krajiny Európskej únie sú zaradené do klastrov pomocou Wardovej metódy. Následne vzťah medzi emisiami CO₂ a hospodárskym rastom je testovaný pomocou Kruskal-Wallis testu.

¹ This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. 1/0245/15 „Transportation planning focused on greenhouse gases emission reduction“.

1.1 Princíp Wardovej metódy

Wardovu metódu z hľadiska publikovaných prác v domácej a zahraničnej literatúre možno považovať za najpoužívanejšiu a zároveň z praktického hľadiska za najvhodnejšiu metódu z hierarchických aglomeratívnych metód vytvárania zhlukov. Táto metóda patrí medzi najprepracovanejšie metódy. Odlišuje sa od iných metód zhlukovania nielen spôsobom tvorby distančnej matice, ale aj spôsobom postupovania pri tvorbe zhlukov. Wardova metóda na prvom kroku využíva pri tvorbe zhlukov najmenšiu štvorcovú Euklidovskú vzdialenosť medzi alternatívami. Ďalšie kroky sú spojené s výberom alternatív, ktoré pri tvorbe ďalšieho zhluku najmenej zväčšia mieru heterogenity zhlukovaných alternatív. V rámci tvorby zhluku sa podľa možnosti minimálne zvýši hodnota rozptylu medzi alternatívami, čím sa zabezpečí tvorba homogénnych zhlukov. Ako miera heterogenity sa využíva variačné kritérium, ktoré sa tiež nazýva chybový súčet štvorcov (rozptyl), resp. vnútrozhluková variabilita. Wardova metóda minimalizácie rozptylu ako špeciálnej cieľovej funkcie bola prvýkrát prezentovaná v práci Ward, J.H.Jr.

Pri tvorbe metodológie zhlukovania Ward vychádzal z existencie menšej skupiny analyzovaných alternatív vzhľadom na veľký počet znakov. Cieľom metódy je vytvorenie zhlukov takým spôsobom, ktorý by minimalizoval straty homogenity asociované s každým zhlukom, a kvantifikoval ich v tvare, ktoré by sa dali ľahko interpretovať. Wardova metóda minimalizácie rozptylu sa realizuje na základe Lance-Williamsovho vzťahu². Pre disjunktné zhluky G_i, G_j a G_k o veľkosti n_i, n_j a n_k objektov platí:

$$d(G_i \cup G_j, G_k) = \frac{n_i + n_k}{n_i + n_j + n_k} d(G_i, G_k) + \frac{n_j + n_k}{n_i + n_j + n_k} d(G_j, G_k) - \frac{n_k}{n_i + n_j + n_k} d(G_i, G_j), \quad (1.1)$$

pričom parametre sú dané vzťahmi

$$\alpha_i = \frac{n_i + n_k}{n_i + n_j + n_k}, \quad \alpha_j = \frac{n_j + n_k}{n_i + n_j + n_k}, \quad \beta = \frac{-n_k}{n_i + n_j + n_k}, \quad \gamma = 0. \quad (1.27)$$

Na základe uvedených vzťahov je možné konštatovať, že kritériom Wardovej metódy spájania zhlukov je prírastok celkového súčtu štvorcov odchýlok pozorovaní vo vnútri skupiny od centroidu zhluku. Pod prírastkom rozumieme súčet štvorcov odchýlok v novo vznikajúcom zhluky, ktorý je zmenšený o súčet štvorcov pôvodných zhlukov. Potom sa spoja zhluky, ktoré majú minimálny rast súčtu štvorcov odchýlok od priemeru zhluku. Táto metóda má tendenciu odstraňovať malé zhluky a vytvára zhluky podobného tvaru a veľkosti.

Algorismus Wardovej metódy

1. vytvorenie distančnej matice (matice nezhody) Z_D medzi alternatívami na základe hodnôt znakov,
2. najnižšia distančná hodnota (z_{d_1}) určuje, ktoré alternatívy budú tvoriť prvý zhluk G_1 ,

² Lance, Williams (1967) a Cormack (1971)

3. pre tieto alternatívy (pre zhluk G_1) sa vypočíta pre každý znak priemerná hodnota,
4. vypočíta sa VZV_1 a výsledky sa zapíšu do dištančnej matice,
5. vyhladá sa najnižšia hodnota v dištančnej matici (z_{d_2}), ktorá určuje, ktorá alternatíva sa bude spájať so zhlukom G_1 ,
6. vypočítajú sa dištančné vzdialenosti medzi zhlukom G_1 a ostatnými alternatívami, a vytvorí sa G_2 na základe vzťahu (4.6); získa sa matica dvojnásobných prírastkov miery heterogenity (označená ako ΔMH),
7. pre minimálnu vzdialenosť medzi G_1 a alternatívami sa vypočíta sa VZV_2 pre nový zhluk G_2 na základe vzťahu $VZV_{G_2} = VZV_{G_1} + 0,5\Delta MH$, čo je chybový súčet štvorcov pre nový zhluk G_2 .

Táto metóda má tendenciu odstraňovať malé zhluky a vytvára zhluky podobného tvaru a veľkosti. Treba poznamenať, že každý zhluk je charakterizovaný inými súradnicami ťažiska.

Výsledky zhľukovania je možné zobrazit' pomocou vývojového dvojrozmerného grafu, ktorý je známy pod názvom dendrogram. Dendrogram je stromový diagram používaný k zobrazeniu vytváraných spojení (zhľukov) na každom kroku analýzy. Základom dendrogramu je celý dátový súbor a každý „list“ je jedna alternatíva dát. Uzly medzi nimi opisujú rozsah, v ktorom sú alternatívy navzájom príbuzné a výška dendrogramu vyjadruje vzdialenosť každého páru alternatív resp. zhľukov. Výsledky zhľukovej analýzy sú získavané prierezom dendrogramu na rôznych úrovniach.

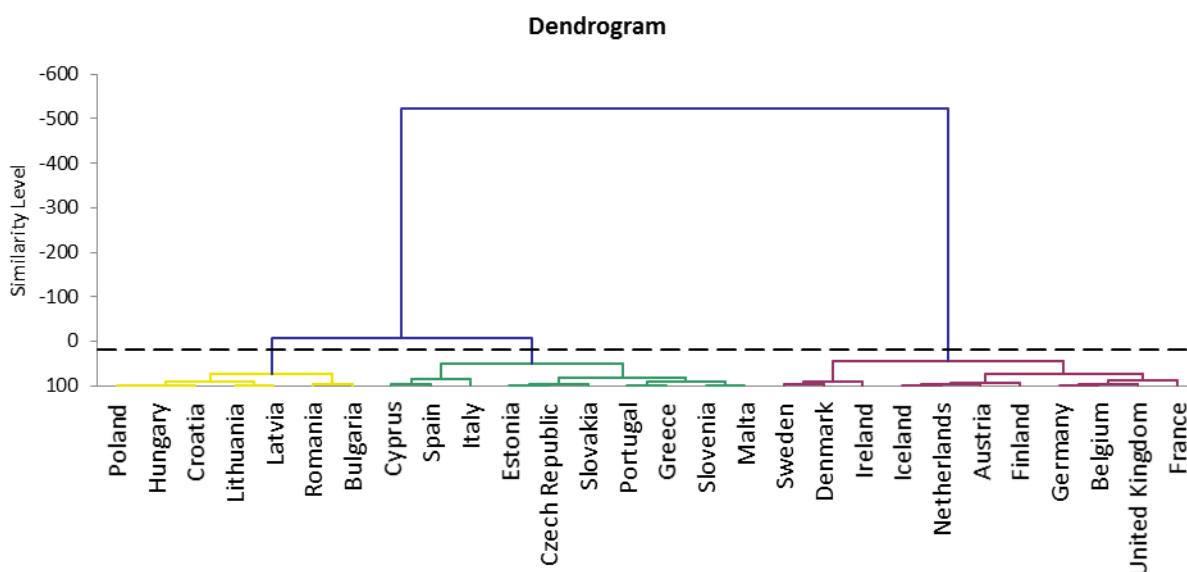
Existujú dva typy dendrogramov. Dendrogram podobnosti alternatív (skúmaných objektov) je štandardný výstup hierarchických zhľukovacích metód, ktorý umožňuje hodnotiť vytvorené zhluky na základe podobnosti analyzovaných alternatív a z ktorého je zjavná štruktúra objektov v zhľukoch.

Z dendrogramu podobnosti znakov je možné vo všeobecnosti zistiť n -tice znakov, ktorých hodnoty sú veľmi podobné a plne spolu koreľujú. Pretože znaky sú v spoločnom zhluku dostatočne podobné a sú vzájomne nahraditeľné, niektoré znaky nie je potrebné kvantifikovať (merať). Dendrogramy je možné využiť aj pri analýze ekonomických kritérií (znakov) ktoré popisujú alternatívy (oblasti), pričom podobné znaky (ukazovatele, vlastnosti), sa môžu substituovať a sú teda ľahko nahraditeľné inými, čím prispievajú do celku svojou vypovedacou schopnosťou.

Výsledky a diskusia.

V tejto časti prezentujeme výsledky klastrovej analýzy získané pomocou Wardovej metódy zobrazené na dendrograme na obrázku číslo 1 a významné rozdiely medzi klastrami z Kruskal-Wallis testu.

Pomocou Wardovej metódy sme rozdelili krajiny Európskej únie do troch skupín - klastrov. V prvej skupine je Poľsko, Maďarsko, Chorvátsko, Litva, Lotyšsko, Rumunsko a Bulharsko. Druhá skupina zahŕňa: Cyprus, Španielsko, Taliansko, Estónsko, Česká republika, Slovensko, Portugalsko, Grécko, Slovinsko a Malta. A do tretej skupiny patrí: Švédsko, Dánsko, Írsko, Island, Holandsko, Rakúsko, Fínsko, Nemecko, Belgicko, Spojené kráľovstvo a Francúzsko.



Obrázok č. 1. Dendrogram 2014
Zdroj: vlastné výpočty

Na overenie štatisticky významných rozdielov medzi dvoma alebo viacerými skupinami je vhodné používať metódy, ktoré umožnia zistiť, aké ukazovatele, majú výrazne inú úroveň v jednom zhluku v porovnaní s ostatnými zhlukmi. Za týmto účelom bol použitý Kruskal-Wallis test. Kruskal-Wallis test je neparametrický test ktorý patrí medzi najčastejšie používané viac výberové poradové testy. Kruskal-Wallis test nepredpokladá normálne rozdelenie v dátach a je oveľa menej citlivý na extrémne hodnoty a preto môže byť použitý, aj keď neboli splnené tieto predpoklady. Hodnota Kruskal-Wallis testu bola vypočítaná pre všetky tri klastre. Test naznačuje, že štatisticky významné rozdiely medzi skupinou jeden, dva a tri na 0,05 hladine významnosti sú v nasledujúcich premenných: emisie CO₂ v tonách na obyvateľa, HDP na obyvateľa, vyprodukovaná energia pre konečnú spotrebu na obyvateľa. Závbery sú založené na základe p-hodnôt, ktoré boli v porovnané s hladinou významnosti ($\alpha = 0,05$).

Záver

S využitím údajov z databázy z Eurostat za rok 2014, sme zistili, že existuje korelácia medzi hospodárskym rozvojom a emisiami CO₂. Môžeme si všimnúť, že klastre efektívne rozdeľujú krajiny EÚ predovšetkým v závislosti na HDP na obyvateľa. Výsledky klastrovej analýzy a testu štatisticky významných rozdielov potvrdzujú, že ekonomický rozvoj je spojený so spotrebou energie a emisiami CO₂.

Použitá literatúra

1. CORMACK, R.M. (1971), "A Review of Classification, "Journal of the Royal Statistical Society, Series A, 134, 321–367.
2. KAYA, Y. (1990) Impact of Carbon Dioxide Emission Control on GNP Growth: Interpretation of Proposed Scenarios. Paper presented to the IPCC Energy and Industry Subgroup, Response Strategies Working Group, Paris.

3. KAYA, Y. YOKOBORI, K. (1997) Environment, Energy, and Economy: strategies for sustainability, United Nations University Press, New York, Paris, pp. 331
4. LANCE, G.N., WILLIAMS W.T. (1967) A general theory of classificatory sorting strategies. I. Hierarchical systems, The Computer Journal 9, 373–380
5. Statistical Office of the European Communities. (2016). EUROSTAT: Regional statistics: Reference guide. Luxembourg: Eurostat.
6. WARD, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of American Statistical Association, 58(301), pp. 236 – 244

Kontaktné údaje

Ing. Marian Reiff, PhD.

prof. Ing. Zlatica Ivaničová, CSc.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 823

email: marian.reiff@euba.sk, zlatica.ivanicova@euba.sk

ANALÝZA EFEKTÍVNOSTI PODNIKOV V ODVETVÍ IT TECHNOLÓGIÍ S POUŽITÍM DEA – SBM MODELU A ROA

EFFICIENCY ANALYSIS OF SELECTED IT COMPANIES BY DEA - SBM MODEL AND ROA

Allan Jose Sequeira Lopez

Abstrakt

DEA (z angl. Data Envelopment Analysis – Analýza dátových obalov) predstavuje dátovo orientovaný prístup na hodnotenie výkonnosti skupiny podnikov označených ako DMU (z angl. Decision Making Unit – Rozhodovacie jednotky), ktorých cieľom je premena rôznych vstupov na rôzne výstupy. DEA sa používa na hodnotenie výkonnosti rôznych druhov podnikov z rôznych odvetví. Pre účely našej analýzy hodnotíme efektívnosť 20 slovenských IT podnikov prostredníctvom DEA – SBM modelu.

***Kľúčové slová:** analýza dátových obalov, efektívnosť, vstupy, výstupy, DEA modely, rozhodovacie jednotky.*

Abstract

Data envelopment analysis (DEA) is a “data-oriented” approach for evaluating the performances of a set of entities called decision-making units (DMUs) which convert multiple inputs into multiple outputs. DEA has been used in evaluating the performances of many different kinds of entities engaged in many different kinds of activities in many different contexts. For the purposes of our paper, the efficiency is measured by a selected model of Data Envelopment Analysis – DEA – SBM model. The analysis is performed on 20 selected Slovak IT companies.

***Keywords:** Data Envelopment Analysis, Efficiency, Inputs, Outputs, DEA models, Decision Making Units.*

1. DEA ANALÝZA

Efektívnosť by sa dala voľne definovať ako pomer medzi výstupmi istej činnosti a vstupmi, ktoré sú na tvorbu výstupov potrebné. V ekonomickej teórii je známy fakt, že ekonomické zdroje - vstupy, sa vo svete nachádzajú len v obmedzenom množstve. Efektívnosť, produkčných jednotiek môže byť všeobecne vyjadrená ako pomer žiaducich výstupov, ktoré hodnotená jednotka produkuje, a vstupov, ktoré pri tejto produkcii spotrebováva. Každá jednotka je teda charakterizovaná súborom vstupných a výstupných numerických údajov. Na analýzu týchto údajov môžeme použiť množstvo rôznych modelov a metód (Jablonský, Dlouhý, 2004, s. 7)

Analýza dátových obalov (z angl. Data Envelopment Analysis – DEA) je relatívne nový dátovo orientovaný prístup využívaný pri meraní a hodnotení efektívnosti množiny produkčných resp. rozhodovacích jednotiek, ktoré transformujú vstupy na výstupy.

Rozhodovacia jednotka (z angl. Decision Making Unit – DMU) je pojem, ktorý zaviedol Charnes a kol. (Charnes a kol., 1978) na popis jednotiek, ktoré budú analyzované pomocou DEA. Týmto pojmom chcel docieľiť, aby sa analýza efektívnosti nevzťahovala len na podniky, ktoré vykazujú zisk, ale na podniky vo všeobecnom ponímaní. DEA vyžaduje pri výpočtoch len obmedzený počet predpokladov. Práve táto skutočnosť umožnila, aby sa DEA analýza využila pri meraní efektívnosti aj v prípadoch, v ktorých zložitosť povahy vzťahov medzi vstupmi a výstupmi DMU zamedzila použitie iných metód (Cooper, Seiford, Zhu, 2011, s. 1-2).

DEA je neparametrická metóda na meranie efektívnosti, ktorá používa nástroje matematického programovania (konkrétne lineárneho). Ako priekopníka v tejto oblasti môžeme označiť Farrella (1959), ktorý nadviazal na práce Debreua (1951) a Koopmansa (1951). Koopmans však definuje technickú efektívnosť striktnejšie. Podľa neho je podnik technicky efektívny iba vtedy, ak sa nachádza na hranici produkčných možností a všetky odchýlkové premenné (prebytok vstupov, resp. nedostatok výstupov) sú rovné nule (Koopmans, 1951). Farrell navrhol novú mieru efektívnosti, ktorá bola založená na výpočte lineárnej konvexnej obalovej krivky a využití dištančných funkcií na meranie vzdialenosti DMU od premietnutého bodu na efektívnej krivke. Farrell používal pojem cenová efektívnosť namiesto alokačnej efektívnosti a termín celková efektívnosť namiesto ekonomickej efektívnosti. (FARRELL, 1957)

Hlavnou výhodou DEA analýzy je skutočnosť, že na rozdiel od pomerových ukazovateľov môžeme hodnotiť efektívnosť DMU pomocou viacerých vstupov a výstupov. Za ďalšiu silnú stránku v porovnaní s kvantifikáciou efektívnosti pomocou pomerových ukazovateľov môžeme považovať, že vstupy a výstupy pri DEA môžu byť vyjadrené nielen v peňažných jednotkách, ale v akýchkoľvek iných merných jednotkách, napríklad v množstve kusov či rozlohe. Z tohto hľadiska je DEA obľúbená najmä v prípadoch, kedy je finančné ohodnotenie vstupov a výstupov zložité (verejný sektor, firmy čeliace odlišným cenovým úrovniam v rôznych krajinách). DEA reprezentuje neparametrický prístup pri odhade produkčnej hranice a to v tom zmysle, že nevyžaduje žiadne predpoklady týkajúce sa jej funkčného tvaru (Luptáčík, 2011).

2. ANALÝZA EFEKTÍVNOSTI VYBRANÝCH IT PODNIKOV

Predmetom našej analýzy je 20 podnikov pôsobiacich v odvetví informačných technológií na Slovensku. Pre účely analýzy sme vybrali DEA - SBM model s predpokladom konštantných výnosov z rozsahu.

SBM DEA model meria efektívnosť pomocou prídavných premenných s_i^+ a s_i^- pričom spĺňa nasledujúce dve dôležité podmienky:

- miera efektívnosti je nezávislá na jednotkách použitých pri vyjadrení vstupov a výstupov,
- miera efektívnosti je monotónne klesajúca funkcia všetkých prídavných premenných priradených vstupom a výstupom, t.j. každé zlepšenie (zhoršenie) akéhokoľ-

vek vstupu aj výstupu vedie k zvýšeniu (zníženiu) miery efektívnosti danej produkčnej jednotky.

SBM DEA model, pre prípad konštantných výnosov z rozsahu v matematickom vyjadrení sa dá skonštruovať nasledovne:

$$\text{Min } \rho = \frac{1 - \frac{1}{m} \sum_{i=1}^m \left(\frac{s_i^-}{x_{iq}} \right)}{1 + \frac{1}{r} \sum_{i=1}^r \left(\frac{s_i^+}{y_{iq}} \right)}$$

za podmienok:

$$\begin{aligned} \sum_{j=1}^n x_{ij} \lambda_j + s_i^- &= x_{iq}, i = 1, 2, \dots, m \\ \sum_{j=1}^n y_{ij} \lambda_j + s_i^+ &= y_{iq}, i = 1, 2, \dots, r \\ \lambda_j &\geq 0, s_i^+ \geq 0, s_i^- \geq 0 \end{aligned}$$

Výraz v čitateli, resp. menovateli účelovej funkcie modelu meria priemernú vzdialenosť vstupov resp. výstupov od efektívnej hranice. Je zrejmé, že pre efektívne jednotky je hodnota účelovej funkcie tohto modelu rovná 1. Je možné dokázať, že pre konštantné výnosy z rozsahu je SBM efektívnosť vždy nižšia alebo rovnaká pri porovnaní s inými DEA modelmi ako napríklad CCR modelom. Zároveň ale platí, že DMU označené ako SBM efektívne sú aj CCR efektívne a naopak .

Produkčná jednotka je teda pri SBM DEA modeli efektívna vtedy, ak má úloha optimálne riešenie ρ^* pričom platí, že s mierou efektívnosti nižšou ako 1 sú produkčné jednotky považované za SBM neefektívne.

Vstupné údaje pre účely analýzy sme čerpali z prílohy týždenníka Trend – TREND TOP V INFORMAČNÝCH TECHNOLOGIÁCH (Trend Top v infotechnológiách, 2014) a z Registra účtovných závierok sa 2015– informačného systému verejnej správy pod správou Ministerstva financií Slovenskej republiky, ktorého prevádzkovateľom je rozpočtová organizácia ministerstva DataCentrum.

Na výpočet údajov pre analýzu môžeme použiť doplnok programu Microsoft Excel s názvom DEA Solver, ktorý je určený na riešenie rôznych úloh matematického programovania. Tento nástroj narába so vstupnými údajmi, ktoré označujeme ako vstupy DMU a výstupy DMU. Doplnok po výpočte zobrazí výstupné informácie na niekoľkých hárkoch dokumentu MS Excel, kde prehľadne zobrazí: (JABLONSKÝ, 2004)

- výsledné poradie podnikov podľa ich efektívnosti,
- efektívnosť vyjadrenú pomocou jedného čísla pre každý podnik osobitne,
- grafické zobrazenie nameraných mier efektívnosti,
- projekciu vstupov a výstupov, ktorá obsahuje také ideálne hodnoty vstupov a výstupov pre každý podnik, pri ktorých by dosahoval maximálnu efektívnosť,

- váhy, ktoré prisúdil jednotlivým vstupom a výstupom, pričom sú zobrazené len tie, ktoré zabezpečujú maximálnu možnú efektívnosť v rámci možností daného podniku,
- sklzy (z angl. slacks), ktoré predstavujú vypočítaný rozdiel medzi reálnou hodnotou a tou, ktorá vychádza ako optimálna, t. j. hodnotu, o ktorú sa má zväčšiť/zredukovať vstup alebo výstup.

Z dôvodu obmedzeného rozsahu uvádzame len vybrané výsledky DEA SBM modelu a hodnotu ROA pre účely porovnania. Na základe výstupov SBM modelu s predpokladom konštantných výnosov z rozsahu sme vypočítali efektívnosť pre 20 vybraných IT podnikov a pre každý z nich sme vypočítali aj ROA za účelom porovnania výsledkov DEA modelu s analýzou na základe tradičných metód – ROA. Cieľom porovnania bolo potvrdiť, že DEA modely sú pre účely analýzy efektívnosti vhodnejšie, keďže umožňujú zahrnúť do analýzy viac premenných – vstupov aj výstupov. Zistili sme, že na základe výsledkov DEA modelov sa viaceré podniky ukázali ako efektívne, pričom hodnoty ukazovateľa ROA boli veľmi nízke. Tento rozdiel je spôsobený faktom, že pre DEA modeloch sme použili viac vstupov a výstupov, pričom ROA berie do úvahy len jeden vstup a výstup.

Poradie	DMU	Skóre	ROA
1	msg life Slovakia, s	1	13%
1	elfa, s.r.o., Košice	1	3%
1	runsys, a.s., Liptov	1	75%
1	Siemens Convergence	1	104%
5	ICZ Slovakia, a.s.,	0,683110	35%
6	I.S.D.D. plus, s.r.o	0,664598	48%
7	Microcomp-Computersy	0,620850	35%
8	BSP SoftwareDistribu	0,505066	27%
9	Sevitech, a.s., Brai	0,444690	22%
10	Wincor Nixodorf, s.r	0,428300	21%
11	DotNet, a.s., Bratis	0,407733	30%
12	Disig, a.s., Bratisl	0,354386	18%
13	CGI Slovakia, s.r.o.	0,243929	11%
14	Itelligence Slovakia	0,209074	8%
15	Interway, a.s., Brat	0,184447	7%
16	Alcasys Slovakia, a.	0,134169	5%
17	Sofos, s.r.o., Brati	0,097085	3%
18	exe, a.s., Bratislav	0,054373	2%
19	FaxCopy, a.s., Brati	0,028808	1%
20	SIT consulting, a.s.	0,005029	-34%

Tabuľka Zhrnutie výsledkov analýz DEA - SBM modelu a ROA pre IT podniky

Na základe úplných výstupov SBM modelu (ktoré tu z dôvodu rozsahu neuvádzame) je možné navrhnúť konkrétne odporúčania na zvýšenie efektívnosti pre tie podniky, ktoré nevykazujú dostatočnú efektívnosť v rámci skupiny porovnávaných podnikov.

Použitá literatúra

- BANKER, R. D. - CHARNES, A.- COOPER, W.W . 1984. *Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis In Management Science*. [online] Pittsburgh-Pennsylvania : Carnegie-Mellon University, 1984.
- FARRELL, M.J. 1957. The Measurement of Productive Efficiency. <http://www.jstor.org/>. [Online] 5. August 1957. [Citace: 20. Január 2015.] <http://www.jstor.org/>.
- CHARNES, A. - COOPER, W. - RHODES, E. 1978. *Measuring the efficiency of decision making units*. miesto neznáme : European Journal of Operations Research, zv. 2, 1978.
- CHARNES, ABRAHAM - COOPER, WILLIAM - LEWIN, ARIE Y.- SEIFORD, LAWRENCE M. 1994. *Data Envelopment Analysis, Theory, Methodology and Applications*. Massachusetts : Kluwer Academic Publishers, 1994. 0-79239480-1.
- JABLONSKÝ, J., DLOUHÝ, M. 2004. *Modely hodnocení efektivity produkčních jednotek*. Praha : Professional Publishing , 2004. 80-86419-49-5.
- TREND TOP V INFOTECHNOLÓGIÁCH*. TREND. 2016. Bratislava : TREND Holding, s.r.o., 2016, Sv. 21. 160152/01.
- KLIEŠTIK, T. 2009. *Kvantifikácia efektivity činností dopravných podnikov pomocou Data Envelopment Analysis*. [online] Žilina : Katedra ekonomiky na Žilinskej Univerzite, 2009.
- KOOPMANS, T. C. 1951. Analysis of production as an efficient combination of activities. *Activity analysis of production and allocation*. 1951.
- LUPTÁČIK, M. 2011. . *Mathematical Optimization and Economic Analysis*. New York : Springer, 2011. 978-0-387-89551-2.
- SEQUEIRA LOPEZ, ALLAN JOSE. 2015. Možnosti využitia DEA metodológie pri analýze efektivity podnikov v odvetví informačných technológií. Bratislava : KOVE FHI – Katedra operačného výskumu a ekonometrie FHI, 2015.
- MINISTERSTVO FINANCI SK. 2011-16. [registeruz.sk](http://www.registeruz.sk). *registeruz.sk*. [Online] Ministerstvo financií Slovenskej republiky, 17. 11 2011-16. <http://www.registeruz.sk>.

Kontaktné údaje

Ing. Allan Jose Sequeira Lopez.
Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky
Dolnozemska 1/b, 852 35 Bratislava

1-MINUTOVÁ INTERVALOVÁ FOREX DATA: DOKÁŽEME SPOČÍTAT HORNÍ MEZ ROZPTYLU?

1-MINUTE INTERVAL FOREIGN EXCHANGE MARKET DATA: CAN WE COMPUTE THE UPPER BOUND OF SAMPLE VARIANCE?

Ondřej Sokol

Abstrakt

Práce se zabývá problematikou výpočtu horní meze výběrového rozptylu v případě, kdy jsou k dispozici pouze minutová vysokofrekvenční data, přičemž se zaměřuje výhradně na reálnou výpočetní náročnost. Obecně je nalezení horní meze výběrového rozptylu ze znalosti pouze intervalových dat NP-těžký problém, ale při splnění určitých podmínek kladených na vstupní intervalová data lze použít některý z efektivních algoritmů. V této práci je zkoumáno, zda podmínky efektivní práce Fersonova algoritmu jsou splněny na minutových intervalových EUR/USD datech z FOREXu. Je ukázáno, že algoritmus po malé úpravě dokáže pracovat nad těmito daty velmi rychle.

Klíčová slova: výběrový rozptyl, intervalová data, vysokofrekvenční data

Abstract

The paper deals with a problem when only interval-valued high-frequency data are available and analyzes the computational complexity of sample variance bounds using real data. Computing the tight upper bound of the sample variance of an interval-valued dataset is known to be NP-hard; however, solution can be found in a reasonable time in many practical applications. The Ferson's algorithm designed for finding the tight upper bound of the sample variance is analyzed. It is tested whether 1-minute interval-valued datasets of EUR/USD from a foreign exchange market fulfill assumptions of the algorithm and if not, then it is analyzed how the computational complexity is affected. Results show that in spite of the theoretical NP-hardness, the analyzed algorithm, after slight adjustment, is indeed useful in the studied case.

Keywords: sample variance, interval, high-frequency data

1 INTRODUCTION

Sample variance is a basic descriptive statistic which is needed for many statistical techniques and methods. We assume that only 1-minute interval data from the foreign exchange market Forex data from 2015 are available. The task is to determine if it is possible to compute the tight upper bound of the sample variance in a reasonable time knowing only minimal and maximal bound of the 1-minute intervals. For example to determine minimal and maximal value of option, exact bounds of sample variance are necessary.

It is known that computation of the exact upper bound of the sample variance is NP-hard problem. However, there are algorithms which, under certain conditions, are able to compute bounds in a reasonable time. We are focused on Ferson's algorithm. The goal of the paper is to determine whether Forex 1-minute data fulfills these conditions.

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

$$\mathbf{x}_1 = [\underline{x}_1, \bar{x}_1], \dots, \mathbf{x}_n = [\underline{x}_n, \bar{x}_n] \quad (1)$$

such that $\underline{x}_i \leq x_i \leq \bar{x}_i$ for every $i = 1, \dots, n$. In our dataset \underline{x}_i and \bar{x}_i correspond to minimum and maximum values in the i -th observation representing 1-minute interval.

Computing statistics over interval data may not be an easy task. While computing the bounds of mean, \underline{E} and \bar{E} , is easy as it is just a sum of lower or upper bounds of intervals, the higher moments are not that simple. The computation of the upper and lower bounds, respectively, reduces to the optimization problems (2):

$$\begin{aligned} \underline{\hat{\sigma}} &= \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 : \underline{x} \leq \mathbf{x} \leq \bar{x} \right\}, \\ \bar{\hat{\sigma}} &= \max_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 : \underline{x} \leq \mathbf{x} \leq \bar{x} \right\}. \end{aligned} \quad (2)$$

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. 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.

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 α -narrowed interval $[x^c - \alpha x^\Delta, x^c + \alpha x^\Delta]$.

Only the computation of the upper bound of sample variance using real high-frequency interval data is studied in this paper. Also, other statistics are often studied, for example Černý and Hladík [3] (t -ratio), Kreinovich [4] and Xiang et al. [1] (entropy), Kreinovich et al. [5] (higher moments) and others. Summaries of approaches to computing various statistics under interval uncertainty can be found in [6], [7].

3 DESCRIPTION OF THE ALGORITHM

This paper uses the Ferson's algorithm which was published in [8]. The algorithm works as follows:

1. We sort all $2n$ endpoints of the narrowed intervals $\frac{1}{n}x_1, \dots, \frac{1}{n}x_n$ 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 \overline{E} and select all intervals $[x_{(k)}, x_{(k+1)}]$, that intersect with $[\underline{E}, \overline{E}]$.
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 $x_{(k+1)} < x_i^C - \frac{1}{n}x_i^A$, then $x_i = \overline{x}_i$,
 - if $x_{(k)} > x_i^C + \frac{1}{n}x_i^A$, 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 $\overline{\hat{\sigma}}$.

The exponential part of the computation complexity depends on the maximal number k of narrowed intervals having a common point. As follows from the Ferson's algorithm the points we deal with are the ones inside of the lower and upper bound of the interval dataset mean. This task can be reformulated in a graph-theoretic way: consider the undirected graph $G_n = (V_n, E_n)$, where n is the number of intervals in a dataset, $V = \{1, \dots, n\}$ and $\{i, j\} \in E_n$ iff $\frac{1}{n}x_i \cap \frac{1}{n}x_j \cap \underline{E} \cap \overline{E} \neq \emptyset$. Let ω_n be the size of a maximal clique of the interval graph G_n . Then the computational complexity of the algorithm is $O(n 2^{\omega_n})$.

In many practical applications ω_n is small enough to allow computing $\overline{\hat{\sigma}}$ in a reasonable time. In paper [9], a hypothesis about the computational complexity of randomly generated interval dataset was formulated.

Hypothesis. Let x_1^C, \dots, x_n^C are sampled from a distribution Φ , the radii x_1^A, \dots, x_n^A are sampled from a nonnegative distribution Ψ and the samples are independent. If Φ is a continuous distribution with finite first and second moments, its density function is limited from above and Ψ has finite first and second moments, then $E(\omega_n) = O(\log n)$.

This hypothesis has not been proven analytically yet but simulations presented in [9] strongly support it. In addition it has been found, based on simulation experiments, that $Var(\omega_n) = O(1)$. The question is whether the hypothesis's conditions are "rare" or "frequent" in common datasets. In this study we focus on 1-minute interval data of the foreign exchange market represented by Forex market.

4 DATA

1-minute data from Forex market from 2015 were obtained from website <http://www.histdata.com/>. The dataset consists of thousands of observations which are defined by date, opening, closing, minimal and maximum values during 1-minute time intervals. The number of the 1-minute intervals for the whole year 2015 is about 362 thousand. The dataset does not contain all data; however, the extent of the dataset is sufficient to estimate expected computational complexity. The huge majority of intervals are not degenerated ($x^\Delta > 0$), only about 4 thousand are degenerated. The data format has 5 full decimal places. So x^c has 5 full decimal places and one which can only attain either 5 or 0. The average min-max boundary is approximately $x_i^\Delta \cong 10^{-4}$ and $x_i^c \cong 1$.

4.1 Different length of samples

The computational complexity of computing the sample variance over hourly, daily and weekly dataset of 1-minute intervals is analyzed. The exponential part of computational complexity depends on the ω_n . Therefore, $E(\omega_{60})$ is estimated in the case of hourly data. To do that, 1000 instances of 60 consecutive intervals with a randomly chosen starting point were selected. From the processing of these instances it follows that $E(\omega_{60}) = 3.036$ with the worst case of ω_{60} estimated as 8. Similar processes were done using the daily and weekly data. The results are following: $E(\omega_{1440}) = 4.802$ with the worst case of 13 and $E(\omega_{10080}) = 6.822$ with the worst case of 22. Therefore, it is possible to compute the sample variance of hourly, daily and weekly data with a standard PC very fast. The estimation of distribution of ω_n for various n is shown in Figure 1.

Note: In the case of monthly or yearly dataset the inequality (3)

$$\frac{1}{n} x_i^\Delta \leq |x_i^c - x_j^c| \text{ iff } \{i \neq j; i, j = 1, \dots, n; x_i^c \neq x_j^c\} \quad (3)$$

is valid for every narrowed interval. Knowing this inequality, the complexity depends on the number of centers of intervals x^c that are in the same point in the interval $[\underline{E}, \overline{E}]$. If they have the same radius x^Δ then the sample variance can be computed in a polynomial time [10]. Otherwise, it is necessary to check every combination which is the most difficult case. Processing of the whole dataset found 29 intersecting narrowed intervals. After checking their radius the number of combinations was reduced to approximately 2^{20} that is necessary to check in order to find the upper bound of the sample variance. This case is computable on a standard PC in a few minutes.

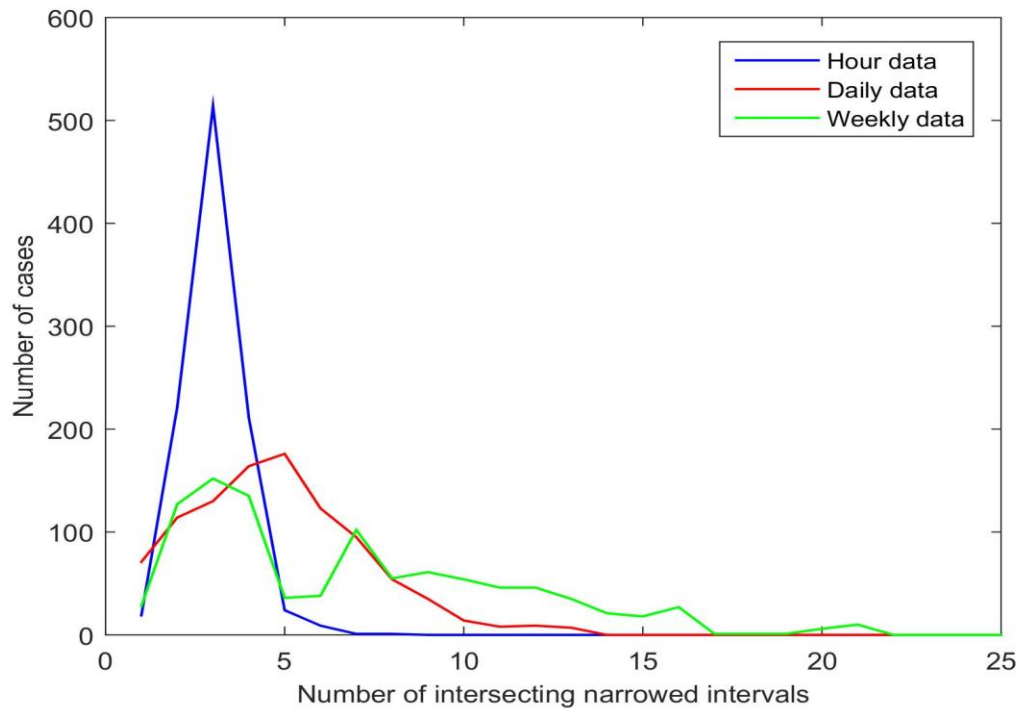


Figure 1: The number of intersecting narrowed intervals based on the length of time series

5 CONCLUSIONS

The analysis of the interval dataset time in case of 1-minute data of the Forex market shows that the upper bound of the sample variance can be computed in a very reasonable time. The structure of the dataset is “friendly” for using the Ferson’s algorithm.

The tight bounds of the sample variance of 1-minute data from Forex market is computable no matter of the length of the time series on the common PC. This finding is the main message of the paper. This result may be useful in analysis of option value on Forex market.

Acknowledgements: This work was supported by IGA F4/54/2015, 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, Vol. 13, No. 6, pp. 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, Vol. 33, No. 2, pp. 108–118.
3. ČERNÝ, Michal and Milan HLADÍK. The Complexity of Computation and Approximation of the t-ratio over One-dimensional Interval Data. *Computational Statistics & Data Analysis*. 2014, Vol. 80, pp. 26–43.
4. KREINOVICH, Vladik. Maximum Entropy and Interval Computations. *Reliable Computing*. 1996, Vol. 2, No. 1, pp. 63–79.

5. KREINOVICH, Vladik, Luc LONGPRÉ, Scott FERSON and Lev GINZBURG. *Computing Higher Central Moments for Interval Data*. B.m.: University of Texas at El Paso. 2004.
6. NGUYEN, Hung T., Vladik KREINOVICH, Berlin WU and Gang XIANG. *Computing Statistics under Interval and Fuzzy Uncertainty*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. Studies in Computational Intelligence.
7. FERSON, Scott, Vladik KREINOVICH, Janos HAJAGOS, William OBERKAMPF and Lev GINZBURG. *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty*. B.m.: Sandia National Laboratories, 2007.
8. FERSON, Scott, Lev GINZBURG, Vladik KREINOVICH, Luc LONGPRÉ and Monica AVILES. Exact Bounds on Finite Populations of Interval Data. *Reliable Computing* [online]. 2005, Vol. 11, No. 3, pp. 207–233.
9. 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.
10. XIANG, Gang. Fast Algorithm for Computing the Upper Endpoint of Sample Variance for Interval Data: Case of Sufficiently Accurate Measurements. *Reliable Computing*. 2006, Vol. 12, No. 1, pp. 59–64.

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

APPLICATION OF DATA ENVELOPMENT ANALYSIS FOR MEASURING OF THE PERFORMANCE OF COUNTRIES AT THE SUMMER OLYMPICS GAMES 2016

APLIKACE ANALÝZY OBALU DAT PRO MĚŘENÍ VÝKONNOSTI ZEMÍ NA LETNÍCH OLYMPIJSKÝCH HRÁCH 2016

Nataliya Soldatyuk

Abstract

This paper presents an application of data envelopment analysis for measuring performance of participating nations at Olympic Games. Proposed DEA model considers two inputs (country's GDP and population size) and three outputs (numbers of gold, silver and bronze medals gained by nations during Olympics). Variable returns to scale are assumed. The results for the last Summer 2016 Olympics in Brazil were analyzed. An efficiency index, radial output expansion and peer groups were computed for each participated nation which gained at least one medal at Olympics.

Keywords: *data envelopment analysis, performance measurement, Summer Olympics 2016*

Abstrakt

Tento příspěvek prezentuje aplikace analýzy obalu dat pro měření výkonnosti zemí na letních olympijských hrách 2016. Navržený DEA model má dva vstupy (HDP a počet obyvatelů země) a tři výstupy (počty získaných zlatých, stříbrných a bronzových medailí). Model předpokládá variabilní výnosy z rozsahu. Pro analýzu byly použité výsledky letních olympijských her 2016 v Brazílii. Index účinnosti a skupiny typu peer byly spočítané pro každou zemi, která získala alespoň jednu medaili na olympijských hrách.

Klíčová slova: *analýza obalů dat, měření výkonnosti, letní olympijské hry 2016*

1 INTRODUCTION

In August 2016 more than 500 thousand foreign travelers arrived to Brazil to watch 10500 athletes vying for 4924 medals on the summer Olympic Games in Rio de Janeiro. Millions of other followers all over the world were following the most important and popular sport event online. Although, the competition is among athletes and not among countries, it is common to present results such as numbers of gold, silver and bronze medals belonging to every country. Each participating nation strives for a better ranking in the Olympics and naturally, the capacity of each nation varies greatly in achieving excellent performance. There are various studies in the literature trying to assess performance of countries on the Olympics. Different techniques, such as regression analysis, data envelopment analysis, neural networks and other were used for these purpose. This study presents results of measuring of national performance on summer Olympics 2016 using model of data envelopment analysis.

2 METHODOLOGY

Data envelopment analysis is a linear-programming-based technique for assessing the relative performance of given units based on inputs and outputs of those units. The input and output data of the existing units (decision making units or DMUs) show what can be achieved with

the technology presently available. However, these DMUs exhibit different degrees of efficiency in transforming inputs into outputs and that is what DEA must assess. One of the advantages of DEA method is that it allows us to measure the efficiency between units with inputs and outputs that are not easily comparable and could be measured in different units (for example, in financial, time or count dimensions). DEA approaches are designed to handle these differences. The aim is to identify which DMUs operate efficiently and therefore belong to the efficiency frontier and which DMUs do not operate efficiently and therefore should make appropriate adjustments in their input and/or output mix in order to improve efficiency.

DEA approach was introduced for the first time in 1978 by Charnes, Cooper and Rhodes. Since that time, it became a popular efficiency evaluation tool and found its widespread application in areas that range from hospitals, sports, universities to supply chains of enterprises, banks and others. There are various examples of application of DEA for evaluation participants' Olympics achievements. Lozano et al. (2002) used two-input and three-output DEA model to measure performance of participating nations at five Summer Olympic Games (1984–2000); Lins et al. (2003) proposed Zero-Sum Gains (ZSG) DEA model to rank participating nations in the Olympic Games, which ensures that total number of gold, silver and bronze medals is constant; Wu et al. (2009) applied the cross efficiency method to assessing nations' performance and used cluster analysis to select appropriate targets for inefficient nations; Zhang et al. (2009) discussed underlying preferences in DEA and proposed DEA models with lexicographic preference to analyze performance of participating nations; finally, Chen et al. (2009) proposed two-stage DEA models, where number of representing athletes is an output of first DEA stage and a input of second DEA stage.

We have performed one-stage DEA analysis of the last Summer 2016 Olympic Games. In our model DMUs correspond to participating nations. Only nations which have gained at least one medal (gold, silver or bronze) were taken into consideration, while nations with zero medals were assumed to have a null performance index. The model is formed of two inputs and three outputs. Similar to prior literature (Lozano et al. 2002), inputs are defined as GDP per capita (in US dollars) measured for prior to Olympics year and population of the participating nation. These two inputs represent nation's economic and demographic power, although other demographic inputs could be included into the model. For instance, some studies consider average population height as it is believed that higher athletes have advantage in some sports. Thus we have considered only two inputs mentioned above. Outputs defined as numbers of gold, silver and bronze medals gained during the Olympics.

Output-orientation is assumed in this paper, since the Olympics achievement is measured with respect to medals won (Li et al. 2008). In addition, as all the DEA literature has stated, such as Lozano et al. (2002), Lins et al. (2003), Wu et al. (2009), the population of participating nations varies greatly, then the variable return to scale (VRS) should be assumed in the model.

Let's assume the following notations:

n	number of units (countries)	0	Index of DMU being assessed
GDP_j	Gross Domestic Product of country j	$u_{GDP,0}$	weight assigned to input GDP
P_j	Population of country j	$u_{P,0}$	weight assigned to input P

NG_j	number of gold medals gained by country j	$v_{NG,0}$	weight assigned to output NG
NS_j	number of silver medals gained by country j	$v_{NS,0}$	weight assigned to output NS
NB_j	number of bronze medals gained by country j	$v_{NB,0}$	weight assigned to output NB
α	number of silver medals that are equivalent to one gold medal	w_0	VRS auxiliary variable
β	number of bronze medals that are equivalent to one silver medal		

The output-oriented LP model with assumption of variable returns to scale could be presented as proposed by Lozano et al. (2002):

Minimize (1)

$$u_{GDP,0}GDP_0 + u_{P,0}P_0 - w_0$$

Under constraints

$$v_{NG,0}NG_0 + u_{NS,0}NS_0 + u_{NB,0}NB_0 = 1$$

$$v_{NG,0}NG_j + u_{NS,0}NS_j + u_{NB,0}NB_j - u_{GDP,0}GDP_j - u_{P,0}P_j + w_0 \leq 0, \quad \forall j, j \in \{1, \dots, n\}$$

$$-v_{NG,0} + \alpha v_{NS,0} \leq 0$$

$$-v_{NS,0} + \beta v_{NB,0} \leq 0$$

$$u_{GDP,0}, u_{P,0}, v_{NG,0}, v_{NS,0}, v_{NB,0} \geq 0, \quad w_0 \text{ free}$$

This model has six continuous variables, five non-negative and one free. The number of constraints is equal to the number of countries plus three. The optimal value of the objective function is the inverse of the radial efficiency of DMU 0. For convenience we have used the corresponding dual model of (1)

Maximize (2)

$$\gamma_0$$

Under constraints

$$\sum_{j=1}^n \lambda_j GDP_j \leq GDP_0,$$

$$\sum_{j=1}^n \lambda_j P_j \leq P_0,$$

$$\sum_{j=1}^n \lambda_j NG_j \geq \gamma_0 NG_0 + \mu_{NG,NS},$$

$$\sum_{j=1}^n \lambda_j NS_j \geq \gamma_0 NS_0 - \alpha \mu_{NG,NS} + \mu_{NS,NB},$$

$$\sum_{j=1}^n \lambda_j NB_j \geq \gamma_0 NB_0 + \beta \mu_{NG,NS},$$

$$\sum_{j=1}^n \lambda_j = 1,$$

$$\lambda_j, \mu_{NG,NS}, \mu_{NS,NB} \geq 0, \quad \gamma_0 \text{ free}$$

The variable γ_0 represents the radial (ie equiproportional) expansion of the three outputs that DMU 0 can achieve with its present GNP and population. The variables $\mu_{NG,NS}$ and $\mu_{NS,NB}$ are the dual variables corresponding to the two weight constraints. Calculations were conducted under assumption that $\alpha = \beta = 1$.

3 RESULTS

In this section the results of applying the described DEA model to the data of the last Summer Olympics are presented. Table 1 shows, for each country that gained at least one medal in Brazil 2016, its radial efficiency θ_0^* , radial output expansion γ_0^* , and peer group together with its associated λ_0^* . The peer group (combined in the proportions defined by their multipliers λ_0^*) are used to compute a virtual participant which is technically efficient and to which the country under evaluation can be compared in order to establish its relative efficiency. Thus, the radial efficiency θ_0^* represents the inverse of the amount by which the number of all three types of medals won by a country could have been simultaneously increased. Value of 1 would mean that it is not possible to increase all three types of medals simultaneously and equiproportional. Alternatively, a value of 0.5 would mean than the number of all three types of medals could be doubled.

DMU	Country	Gold	Silver	Bronze	θ_0^*	γ_0^*	Benchmark(Lambda)	Times as a benchmark
1	United States	46	37	38	1.00	1.00	1(1.00)	8
2	Great Britain	27	23	17	1.00	1.00	2(1.00)	26
3	China	26	18	26	1.00	1.00	3(1.00)	7
4	Russia	19	18	19	1.00	1.00	4(1.00)	23
5	Germany	17	10	15	0.82	1.21	1(0.116); 2(0.525); 20(0.26); 39(0.09)	0
6	Japan	12	8	21	1.00	1.00	1(0.35); 20(0.248); 39(0.402)	0
7	France	10	18	14	0.87	1.15	2(0.795); 4(0.095); 39(0.109)	0
8	South Korea	9	3	9	0.61	1.64	1(0.047); 2(0.438); 20(0.068); 39(0.446)	0
9	Italy	8	12	8	0.66	1.53	2(0.626); 4(0.119); 39(0.255)	0
10	Australia	8	11	10	0.92	1.09	2(0.282); 19(0.221); 39(0.497)	0
11	Netherlands	8	7	4	0.79	1.27	2(0.222); 16(0.534); 19(0.244)	0
12	Hungary	8	3	4	0.99	1.01	2(0.109); 16(0.787); 19(0.089); 39(0.014)	0
13	Brazil	7	6	6	0.38	2.66	3(0.054); 4(0.887); 15(0.06)	0
14	Spain	7	4	6	0.44	2.30	2(0.527); 4(0.061); 16(0.057); 39(0.355)	0
15	Kenya	6	6	1	1.00	1.00	15(1.00)	11
16	Jamaica	6	3	2	1.00	1.00	16(1.00)	20
17	Croatia	5	3	2	0.79	1.26	2(0.021); 16(0.91); 19(0.068)	0
18	Cuba	5	2	4	0.82	1.22	2(0.037); 4(0.032); 16(0.709); 39(0.355)	0
19	New Zealand	4	9	5	1.00	1.00	19(1.00)	14
20	Canada	4	3	15	1.00	1.00	20(1.00)	10
21	Uzbekistan	4	2	7	1.00	1.00	21(1.00)	8
22	Kazakhstan	3	5	9	0.82	1.22	1(0.007); 2(0.086); 20(0.036); 39(0.87)	0
23	Colombia	3	2	3	0.34	2.93	4(0.293); 16(0.445); 21(0.097); 39(0.165)	0
24	Switzerland	3	2	2	0.45	2.24	2(0.076); 16(0.48); 19(0.444)	0
25	Iran	3	1	4	0.32	3.12	4(0.435); 21(0.476); 39(0.089)	0
26	Greece	3	1	2	0.40	2.52	2(0.114); 16(0.559); 19(0.259); 39(0.067)	0
27	Argentina	3	1	0	0.24	4.23	2(0.194); 4(0.201); 16(0.604)	0
28	Denmark	2	6	7	1.00	1.00	28(1.00)	7
29	Sweden	2	6	3	0.59	1.70	2(0.086); 19(0.914)	0
30	South Africa	2	6	2	0.61	1.65	4(0.283); 15(0.197); 39(0.52)	0
31	Ukraine	2	5	4	0.88	1.13	4(0.128); 15(0.162); 34(0.645); 39(0.065)	0
32	Serbia	2	4	2	0.75	1.33	16(0.336); 39(0.589); 42(0.05); 68(0.023)	0
33	Poland	2	3	6	0.47	2.12	1(0.085); 20(0.072); 39(0.843)	0
34	North Korea	2	3	2	1.00	1.00	34(1.00)	3

35	Belgium	2	2	2	0.31	3.20	2(0.107); 19(0.87); 39(0.023)	0
36	Thailand	2	2	2	0.20	4.95	4(0.37); 15(0.19); 16(0.21); 21(0.05); 39(0.1)	0
37	Slovak Republic	2	2	0	0.36	2.80	2(0.026); 16(0.605); 19(0.3); 39(0.069)	0
38	Georgia	2	1	4	1.00	1.00	38(1.00)	6
39	Azerbaijan	1	7	10	1.00	1.00	39(1.00)	38
40	Belarus	1	4	4	0.59	1.68	2(0.007); 4(0); 16(0.098); 39(0.894)	0
41	Turkey	1	3	4	0.27	3.71	1(0.044); 4(0.403); 39(0.553)	0
42	Armenia	1	3	0	1.00	1.00	42(1.00); (0)	2
43	Czech Republic	1	2	7	0.70	1.44	2(0.015); 19(0.012); 39(0.973)	0
44	Ethiopia	1	2	5	1.00	1.00	44(1.00); (0)	2
45	Slovenia	1	2	1	0.49	2.06	16(0.092); 19(0.37); 51(0.025), 69(0.512)	0
46	Indonesia	1	2	0	0.21	4.67	3(0.151); 4(0.128); 15(0.721)	0
47	Romania	1	1	3	0.27	3.67	1(0.016); 2(0.069); 4(0.009); 39(0.906)	0
48	Bahrain	1	1	0	0.42	2.39	16(0.339); 19(0.089); 69(0.572)	0
49	Vietnam	1	1	0	0.14	7.20	3(0.031); 4(0.069); 15(0.9)	0
50	Taiwan	1	0	2	0.17	6.00	2(0.174); 20(0.154); 39(0.672); 39(0.662)	0
51	Bahamas	1	0	1	1.00	1.00	51(1.00)	5
52	Ivory Coast	1	0	1	0.34	2.96	15(0.003); 16(0.141); 21(0.194); 34(0.662)	0
53	Fiji	1	0	0	1.00	1.00	53(1.00); (0)	1
54	Jordan	1	0	0	0.16	6.32	4(0.025); 15(0.078); 16(0.897)	0
55	Kosovo	1	0	0	1.00	1.00	55(1.00)	0
56	Puerto Rico	1	0	0	0.16	6.25	2(0.012); 16(0.988)	0
57	Singapore	1	0	0	0.14	6.97	2(0.046); 16(0.954)	0
58	Tajikistan	1	0	0	1.00	1.00	58(1.00)	0
59	Malaysia	0	4	1	0.40	2.48	2(0.1); 4(0.12); 39(0.779)	0
60	Mexico	0	3	2	0.18	5.41	2(0.016); 4(0.815); 39(0.169)	0
61	Algeria	0	2	0	0.26	3.91	4(0.112); 15(0.41); 39(0.479)	0
63	Lithuania	0	1	3	0.90	1.11	28(0.091); 38(0.596); 51(0.313)	0
64	Bulgaria	0	1	2	0.27	3.72	28(0.043); 38(0.404); 39(0.553)	0
65	Venezuela	0	1	2	0.16	6.08	1(0.059); 20(0.101); 39(0.84)	0
66	India	0	1	1	0.20	5.11	3(0.087); 15(0.43); 44(0.483)	0
67	Mongolia	0	1	1	0.51	1.98	16(0.158); 38(0.415); 42(0.362); 53(0.064)	0
68	Burundi	0	1	0	1.00	1.00	68(1.00)	3
69	Grenada	0	1	0	1.00	1.00	69(1.00)	4
70	Niger	0	1	0	0.65	1.54	34(0.27); 68(0.73)	0
71	Philippines	0	1	0	0.12	8.43	3(0.029); 4(0.173); 15(0.797)	0
72	Qatar	0	1	0	0.20	4.97	19(0.496); 69(0.504)	0
73	Norway	0	0	4	0.62	1.61	28(0.905); 51(0.095)	0
74	Egypt	0	0	3	0.31	3.27	3(0.029); 4(0.19); 21(0.781)	0
75	Tunisia	0	0	3	0.42	2.36	21(0.059); 39(0.668); 68(0.273)	0
76	Israel	0	0	2	0.22	4.57	28(0.289); 39(0.711)	0
77	Austria	0	0	1	0.11	9.25	28(0.25); 39(0.75)	0
78	Dominican R.	0	0	1	0.10	10.06	20(0.011); 39(0.989)	0
79	Estonia	0	0	1	0.55	1.83	38(0.282); 51(0.702); 69(0.016)	0
80	Finland	0	0	1	0.15	6.55	28(0.79); 38(0.18); 39(0.03)	0
81	Moldova	0	0	1	1.00	1.00	81(1.00)	0
82	Morocco	0	0	1	0.13	7.93	4(0.046); 21(0.829); 39(0.125)	0
83	Nigeria	0	0	1	0.11	8.93	3(0.111); 21(0.801); 44(0.088)	0
84	Portugal	0	0	1	0.10	10.11	20(0.021); 39(0.979)	0
85	Trinidad	0	0	1	0.52	1.93	28(0.052); 38(0.207); 51(0.741)	0
86	UAE	0	0	1	0.10	10.02	20(0.003); 39(0.997)	0

Table 1. Results of DEA efficiencies measurement.

It can be seen that out of the 86 participants that won medals only 22 are efficient: United States, Great Britain, China, Russian Federation, Kenya, Jamaica, New Zealand, Canada, Uzbekistan, Denmark, North Korea, Georgia, Azerbaijan, Armenia, Ethiopia, Bahamas, Fiji, Kosovo, Tajikistan Burundi, Grenada and Moldova. Of these, the ones which appear more frequently as peer (ie benchmark) are Azerbaijan (38 times), Great Britain (26 times), Russia

(23 times) and Jamaica (20 times). It is not surprising that, China and USA are efficient, but have a low peer count number, as they have very large population in comparison with other countries. Azerbaijan, in other hand, represents opposite, but more common situation: a small country both in population and domestic product. The peer count number can be considered a measure of the extent to which the performance of an efficient country can be a useful reference for the non-efficient. For some countries such as North Korea and Moldova, resulting 100% efficiency could be explained by very low value of GDP (for North Korea, GDP per capita is 583 USD in 2014). Interestingly, that, according to DEA results, performance of Brazil is only 38%, which contradicts with hypothesis proposed by Lozano (2002) saying that organizing countries usually have high performance.

4 SUMMARY AND CONCLUSIONS

This paper has presented an application of a simple DEA model for assessing performance of participating nations at the Olympic Games 2016. Proposed DEA model considers two inputs (country's GDP and population size) and three outputs (numbers of gold, silver and bronze medals gained by nations during the Olympics). The model assumes variable returns to scale. Data from the Summer 2016 Olympics in Brazil were analysed and performance of each participating nation was estimated. For some participants, high Olympic achievements corresponds to high measured performance (for instance, USA, China and Great Britain), but for other participating nations results show lower performance (for example, Germany, Italy, Brazil) and thus these participants could have gained higher number of medals.

5 ACKNOWLEDGEMENTS

This paper is supported by the research grant IGA VŠE/IG403025/54/2015.

6 REFERENCES

1. ANOUZE, A., OSMAN, I., EMROUZNEJAD, A. (2014) Handbook of Research on Strategic Performance Management and Measurement Using Data Envelopment Analysis, IGI Global., ISBN: 978-1-4666-4474-8.
2. BOUSSOFIANE A., DYSON R.G., THANASSOULIS E. (1991) Applied data envelopment analysis. European Journal of Operational Research, vol. 52, p.1-15.
3. CHARNES A., COOPER W.W., RHODES E. (1978) Measuring the efficiency of decision-making units. European Journal of Operational Research, vol. 2, p.249-444
4. CHURILOV, L., FLITMAN, A. (2006). Towards fair ranking of Olympics achievements: The case of Sydney 2000. Computers & Operations Research, 33(7), 2057–2082.
5. COOPER W., SEIFORD M., TONE K. (2007). Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Second edition. ISBN 978-0-387-45283-8.
6. CHEN, Y., COOK, W. D., LI, N., ZHU, J. (2009). Additive efficiency decomposition in two-stage DEA. European Journal of Operational Research, 196(3), 1170–1176.
7. LINS, M. P. E., GOMES, E. G., SOARES DE MELLO, J. C. C., SOARES DE MELLO, A. J. R. (2003). Olympic ranking based on a zero sum gains DEA model. European Journal of Operational Research, 148(2), 312–322.

8. LOZANO, S., VILLA, G., GUERRERO, F., CORTÉS, P. (2002). Measuring the performance of nations at the Summer Olympics using data envelopment analysis. *Journal of the Operational Research Society*, 53(5), 501–511.
9. WU, J., LIANG, L., CHEN, Y. (2009). DEA game cross-efficiency approach to Olympic rankings. *Omega*, 37(4), 909–918.
10. ZHANG, D., LI, X., MENG, W., LIU, W. (2009). Measuring the performance of nations at the Olympic Games using DEA models with different preferences & star. *Journal of the Operational Research Society*, 60(7), 983–990.
11. EMROUZNEJAD, A., PODINOVSKI, V. (2004). *Data Envelopment Analysis and Performance Management*, Warwick Print, Coventry, UK, ISBN: 0 902683 73 X

Kontaktní údaje

Mgr. Nataliya Soldatyuk

Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky, Katedra ekonometrie
nám. W. Churchilla 4, 130 67 Praha 3

Tel: (420) 224 095 423

email: xsoln900@vse.cz

NEZAMESTNANOSŤ A HOSPODÁRSKE CYKLY¹

UNEMPLOYMENT AND BUSINESS CYCLE

Kvetoslava Surmanová

Abstrakt

V makroekonomických modeloch sa zvyčajne môžeme stretnúť s problémom, ktorým sú fluktuácie v zamestnanosti, resp. nezamestnanosti v súvislosti od fázy hospodárskeho cyklu. Zamestnanosť a nezamestnanosť zvyčajne veľmi kolíšu počas hospodárskeho cyklu. Cieľom príspevku odhadnúť rovnovážnu mieru nezamestnanosti a porovnať ju s cyklickým vývojom národného hospodárstva Slovenskej republiky.

Kľúčové slová: nezamestnanosť, NAIRU, ekonomický cyklus, Hodrick – Prescottov filter

Abstract

The macroeconomic models usually can meet the problems that are fluctuations in employment, respectively the unemployment in the context of the business cycle. Employment and unemployment usually fluctuate widely throughout the economic cycle. Aim of this paper to estimate the equilibrium unemployment rate and compare it with the cyclical development of the national economy of the Slovak Republic.

Keywords: unemployment, NAIRU, business cycle, Hodrick – Prescott filter

ÚVOD

Hospodársky rast je možné vyjadriť prostredníctvom rastu reálneho hrubého domáceho produktu (HDP). Podľa fázy hospodárskeho cyklu dochádza k odlišnému vývoju makroekonomických ukazovateľov vrátane dopadov na trh práce. Vznik hospodárskych cyklov je najčastejšie spájaný s existenciou šokov v ekonomike. Týmto šokom je zväčša krátkodobý ekonomický šok (napr. fiškálny, peňažný, atď.), ktorý zapríčiňuje dočasné odchýlky skutočného agregátneho výstupu od dlhodobej trajektórie vývoja potenciálneho produktu. Tradičný pohľad predpokladá, že trend vývoja potenciálneho produktu je viac menej hladký a produkčné možnosti ekonomiky rastú stabilným tempom bez väčších výkyvov. Zatiaľ čo teórie hospodárskeho cyklu skúmajú príčiny opakovaného vzniku recesií a inflačných medzier výstupu ako prechodných ekonomických javov, teória rastu skúma príčiny trvale pozorovaného zvyšovania hranice produkčných možností.

Zachytiť vzťah medzi ekonomickým rastom a nezamestnanosťou je možné Okunovým zákonom. Tento nepriamo úmerný vzťah má niekoľko variantov a spôsobov definovania uvedenej závislosti. Iným spôsobom je gapová verzia, ktorá vyžaduje odhad potenciálneho produktu a prirodzenej miery nezamestnanosti. V nasledujúcich častiach príspevku sa zameriame na hospodársky cyklus a NAIRU (miera nezamestnanosti, ktorá nezrýchľuje infláciu).

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

1. HOSPODÁRSKY CYKLUS A NAIRU

V literatúre sa môžeme stretnúť s rôznymi pohľadmi na hospodársky cyklus. Podľa Hodricka a Prescottta sú pozorované výkyvy reálneho výstupu skôr výsledkom pôsobenia ekonomických šokov, z ktorých každý trvale ovplyvňuje produkčné možnosti ekonomiky (napr. náhodné zmeny v produktivite výrobných faktorov) (Hodrick, Prescott, 1997). Kompromisom medzi týmito prístupmi je koncepcia hospodárskeho cyklu ako prejavu spoločného pôsobenia krátkodobých a dlhodobých šokov, čo je označené ako teória náhodnej prechádzky (random walk with drift). Teória v sebe zahŕňa tak tradičnú teóriu hospodárskeho cyklu ako aj teóriu rastu. Podľa teórie náhodnej prechádzky kolíše reálny výstup ekonomiky v čase pod vplyvom dočasne pôsobiacich dopytových šokov a trvale pôsobiacich ponukových šokov. Dočasné šoky pritom kolíšu okolo náhodného (stochastického) trendu. V dlhobehnej perspektíve tak výstup rastie z toho dôvodu, že pozitívne ekonomické šoky sú výraznejšie a prevládajú nad negatívnymi šokmi (Barry, 1996).

Podľa fázy hospodárskeho cyklu dochádza k odlišnému vývoju základných makroekonomických ukazovateľov vrátane dopadov na trh práce, čo najcitlivejšie vníma práve obyvateľstvo. Ide hlavne o ukazovatele ako sú počet zamestnaných, resp. nezamestnaných v ekonomike, výška priemernej mzdy v hospodárstve a vybraných sektoroch. Štát a samotný podnikatelia sledujú vývoj trhu práce rovnako intenzívne, pričom dôraz kladú aj na ďalší dôležitý faktor, ktorým je produktivita práce v danej krajine. Jednotlivé ukazovatele trhu práce sa však nevyvíjajú rovnako v čase recesie a v období expanzie.

Vzťahy vybraných ukazovateľov ekonomického vývoja krajín k hospodárskemu cyklu skúmajú vo svete inštitúcie ako OECD, Eurostat, či Conference Board. Tieto inštitúcie vytvárajú kompozitné indikátory (CLI – composite leading indicators) zamerané na krátkodobé prognózovanie cyklického vývoja ekonomiky. Okrem skupiny tzv. predstihových indikátorov skúmajú aj súbežné a oneskorené indikátory (Czesaný, Jeřábková, 2009). Do týchto troch skupín indikátorov tieto inštitúcie zaraďujú aj ukazovatele trhu práce.

Existuje však úzka spojitosť medzi odchýlkami skutočnej nezamestnanosti od jej rovnovážnej úrovne danej NAIRU a produkčnou medzerou, ktorá reprezentuje cyklický vývoj celkového produktu okolo svojej potenciálnej úrovne. NAIRU je dôležitý ukazovateľ pre hospodársku politiku, pomáha vytvárať obraz o cyklickej pozícii ekonomiky, čo môže byť východisko pri tvorbe a nastavení nástrojov hospodárskej politiky. Stručný prehľad economickej teórie vŕahujúcej sa k hypotéze prirodzenej miery nezamestnanosti a NAIRU možno nájsť v Nemeč, Vašíček (2007).

2. HODRICKOV - PRESCOTTOV FILTER (HP FILTER)

HP filter patrí medzi štatistické metódy. Jeho veľkou výhodou je nenáročnosť na vstupné dáta a je ho možné aplikovať na akýkoľvek časový rad. Nevýhodou je, že pri aplikácii je potrebné zadať vstupný parameter λ , ktorý optimalizuje vyhladenie trendu, nakoľko HP filter vyhladzuje sledovaný časový rad, t. j. očisťuje ho od trendovej zložky. Metóda bola po prvý krát použitá Hodrickom a Prescottom v 1997, pričom analyzovali povojnové hospodárske cykly USA. Ide o dekompozíciu skúmaného časového radu (v logaritmickej vyjadrení) na trendovú y_t^p a cyklickú zložku c_t :

$$y_t = y_t^p + c_t. \quad (1)$$

Proces HP filtra je určený nasledujúcou minimalizačnou funkciou:

$$\{y_t^p\}_{t=0}^{n+1} = \arg \min \sum_{t=1}^n \left\{ (y_t - y_t^p)^2 + \lambda \left[(y_{t+1}^p - y_t^p) - (y_t^p - y_{t-1}^p) \right]^2 \right\}. \quad (3)$$

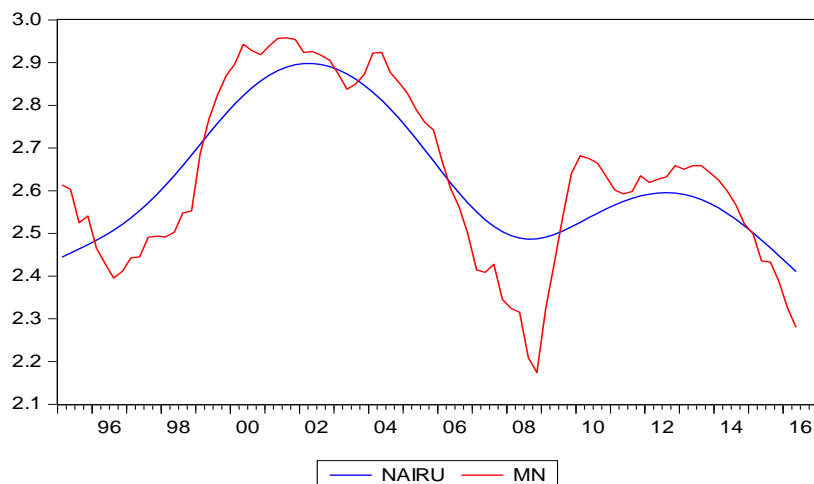
Kde λ^2 je kladné číslo, parameter vyhľadania, alebo sa môžeme nazývať i parameter penalizácie.

3. EMPIRICKÉ VÝSLEDKY

Na analýzu boli použité štvrťročné údaje databázy Slovstat Štatistického úradu SR, za obdobie rokov 1995 – 2016. Makroekonomické ukazovatele HDP (definovaný v mil. Eur v stálych cenách) a miera nezamestnanosti (MN, definovaná v percentách) boli použité v logaritmickom vyjadrení a oba časové rady boli sezónne očistené procedúrou Cenzus X-12. Pri analýze bol využitý softvér Eviews.

Nakoľko bola analýza realizovaná na štvrťročných údajoch, pri aplikácii HP filtra bol parameter vyhľadania, $\lambda=1600$. Následne boli odhadnuté hodnoty potenciálneho HDP a NAIRU. Na obrázkoch 1 a 2 sú znázornené grafické priebehy oboch makroekonomických veličín, ich potenciálnych, rovnovážnych hodnôt aj s ich skutočnými hodnotami.

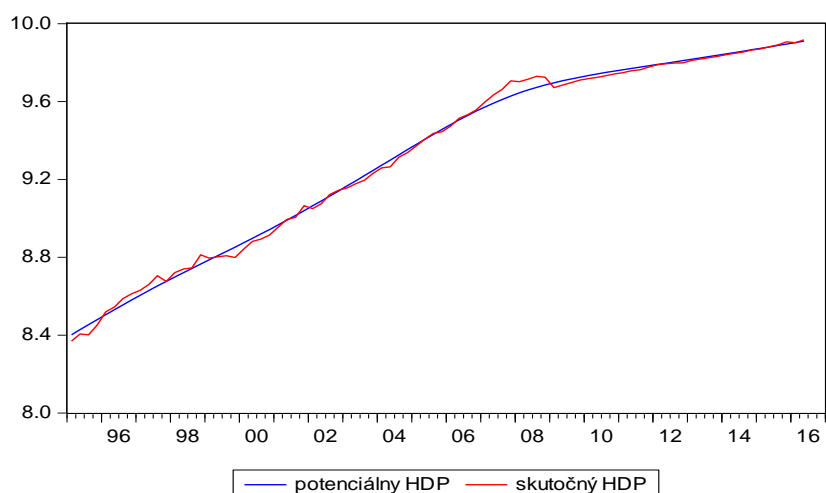
Obrázok 1 Skutočná miera nezamestnanosti a NAIRU



Zdroj: vlastné výpočty

² Hodrick a Prescott navrhli použiť hodnotu $\lambda = 100$ (pre ročné dáta), $\lambda = 1600$ (pre štvrťročné dáta) a $\lambda = 14400$ (pre mesačné dáta).

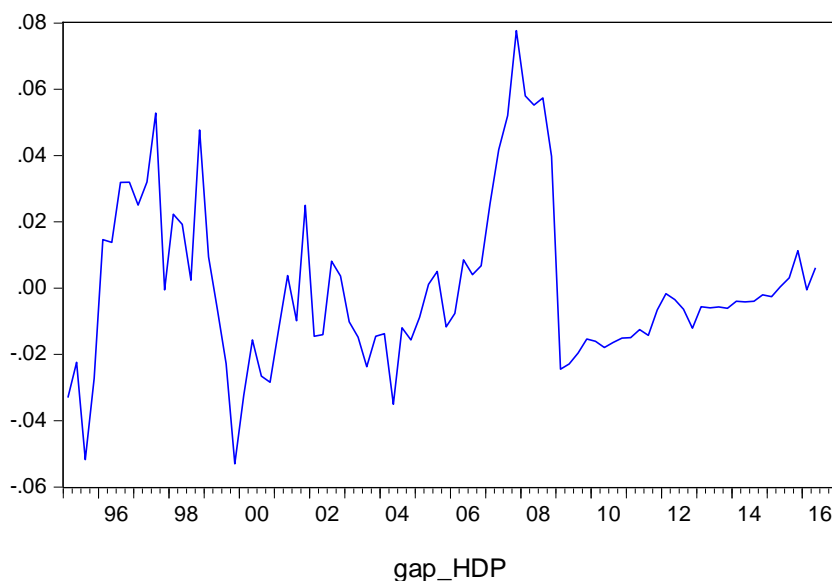
Obrázok 2 Skutočný HDP a potenciálny HDP



Zdroj: vlastné výpočty

Na základe našich výpočtov môžeme na sledovanom období za obdobie rokov 1995Q1 – 2016Q2 pozorovať viacero cyklov, čo je viditeľné aj z obrázka 3, kde je znázornená produkčná medzera (gap_HDP).

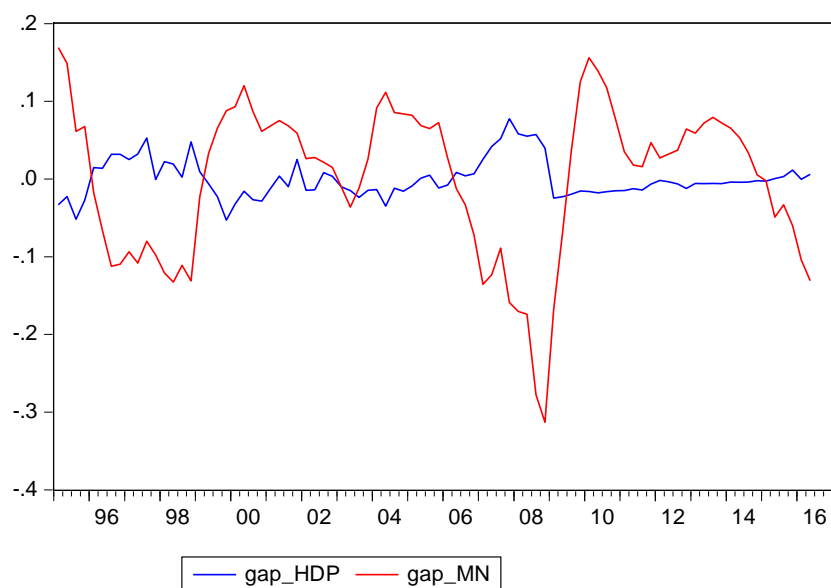
Obrázok 3 Produkčná medzera (gap_HDP)



Zdroj: vlastné výpočty

Z priebehu produkčnej medzery môžeme pozorovať tri, resp. štyri hospodárske cykly, ktorými prešla ekonomika Slovenska za sledované obdobie. Prvý hospodársky cyklus, ktorý je viditeľný končí v roku 1999. Druhý cyklus začal v druhej polovici roka 2000, pričom vrchol dosiahol v roku 2001. Ďalší pozorovateľný cyklus začína v roku 2004 a vrchol dosiahol v 2007. Rok 2009 je možné považovať za jeho konečnú fázu, pretože do vývoja ekonomiky zasiahla finančná kríza.

Obrázok 4 Produkčná medzera (gap_HDP)
a medzera v nezamestnanosti (gap_MN)



Zdroj: vlastné výpočty

Na obrázku 4 sú zachytené obe medzery a z ich priebehu je zrejmé, že zmena, ktorá nastala na strane outputu sa nie vždy prejavila na zmene v nezamestnanosti. Špecifickým sa zdá byť obdobie po roku 2009, kedy je priebeh outputu bez výrazných výkyvov, čo ale nemôžeme tvrdiť o priebehu medzery v nezamestnanosti.

ZÁVER

Odhadom rovnovážnej miery nezamestnanosti a následným porovnaním s cyklickým vývojom národného hospodárstva Slovenskej republiky môžeme konštatovať, že relatívne vysoká úroveň NAIRU na Slovensku naznačuje, že dlhodobo vyššia miera nezamestnanosti nebola výsledkom pretrvávajúcej nerovnováhy v ekonomike, musela by sa totiž v zmysle koncepcie NAIRU prejavit' v rovnako dlhotrvajúcej zápornej produkčnej medzere a nízkej inflácii. Príčiny vysokej nezamestnanosti majú teda zrejme skôr štrukturálny ako cyklický charakter.

Uvedené tvrdenie podporuje aj štruktúra nezamestnaných z hľadiska dĺžky trvania nezamestnanosti. V porovnaní s EÚ a eurozónou je na Slovensku výrazne vyšší podiel dlhodobo nezamestnaných (viac ako rok). Podľa údajov EUROSTAT-u (Labour Force Surveys) bola dlhodobá miera nezamestnanosti na Slovensku najvyššia v rokoch 2008 až 2011 (keď za toto obdobie vzrástla zo 6,7 % na 9,3 %). Až v roku 2012 bolo Slovensko predstihnuté ďalšími krajinami EÚ, ktoré sa dostali do hlbokkej ekonomickej krízy a to Gréckom (14,5 %) a Španielskom (11,0 %). Do roku 2013 na Slovensku však ďalej vzrástla miera dlhodobej nezamestnanosti na 10%. Pomerne veľká časť nezamestnaných ostáva v evidencii úradov práce aj dlhšie ako dva roky a viac.

Navyše je situácia na trhu práce veľmi regionálne diferencovaná. Dlhodobo nezamestnaní nemôžu v najzaostalejších regiónoch Slovenska nájsť na voľnom trhu práce pracovné miesta, sú odkázaní na krátke, príležitostné, zle platené práce a veľakrát v šedej ekonomike. Dlhodobá nezamestnanosť má navyše odozvu aj na fyzickom a psychickom zdraví obyvateľstva, na ich životných hodnotách, ovplyvňuje reprodukčný proces, politické postoje

atď. To sú dôvody, ktoré je nevyhnutné riešiť a malo by byť snahou každej trhovej ekonomiky prostredníctvom fungujúceho systému prevencie im predchádzať.

Použitá literatúra

1. BARRY D. H. 1996. Random walks and random environments, Oxford: Oxford University Press, 1996. ISBN 0-19-853789-1
2. CZESANÝ, S. - JEŘÁBKOVÁ, Z. 2009. Metóda konstrukce kompozitních indikátorů hospodářského cyklu pro českou ekonomiku. In: Statistika, 2009. č.1, s. 21-31. Dostupné na: <http://panda.hyperlink.cz/cestapdf/pdf09c1/czesany.pdf>
3. DUGASOVÁ, B. – FRENÁKOVÁ, A. 2011. Hospodársky cyklus Slovenska s dôrazom na trh práce. In: International Scientific Conference YOUNG SCIENTISTS 2011, dostupné na: http://www3.ekf.tuke.sk/mladivedci2011/herlany_zbornik2011/dugasova_barbora.pdf
4. GYLÁNIK, M. – HUČEK, J. 2009. Odhad NAIRU v slovenskej ekonomike. In: Biatec, roč. 17, 2009, č. 12, str. 13-18
5. HODRIC, J. – PRESCOTT, M. 1997. Postwar U.S. Business cycles: An Empirical Investigation. In: Journal of Money credit and banking 29, 1997. num. 1, pp.1-16.
6. NEMEC, D. – VAŠÍČEK, O. 2007. Alternative Estimates of NAIRU for the Czech Economy: Implications for Economic Growth and Stability. Masaryk University Brno.
7. SURMANOVÁ, K. – FURKOVÁ, A. 2011. Methods for output gap detection. In *Mathematical methods in Economics 2011: proceedings - part I, II*, Praha: Professional Publishing, 2011. ISBN 978-80-7431-058-4, s. [275-280].
8. SZOMOLÁNYI, K. - LUKÁČIK, M. - LUKÁČIKOVÁ, A. 2014. Small open economy real business cycle model of the Slovak economy. In MME 2014 Mathematical methods in economics. international conference. Olomouc, Czech Republic, Olomouc : Palacký University, 2014. ISBN 978-80-244-4209-9, s. 1015-1020 CD-ROM.
9. <http://ec.europa.eu/eurostat/data/database>
10. www.statistics.sk

Kontaktné údaje

Ing. Kvetoslava Surmanová, PhD.

Ekonomická univerzita v Bratislave, Fakulta hospodárskej informatiky

Dolnozemska 1/b, 852 35 Bratislava

Tel: (421 2) 67 295 823

email: kvetoslava.surmaova@euba.sk

SÚ PODIELY NA PRODUKCII DLHODOBO KONŠTANTNÉ?¹

ARE FACTOR SHARES CONSTANT IN THE LONG-RUN?

Karol Szomolányi

Abstrakt

V práci sa zameriame na ekonometrickú kointegračnú analýzu podmienky prvého rádu firmy pre prácu. Využitím agregovaných amerických ročných údajov z Jorgensonových sektorových tabuliek testujeme hypotézu dlhodobej konštantnosti elasticity substitúcie vstupov, ktorá je predpokladom neoklasickej teórie rastu. Okrem toho, že nezamietneme danú hypotézu, odhadneme krátkodobú elasticitu substitúcie vstupov. Jej hodnota sa pohybuje okolo 0,5.

KLúčové slová: *podiely na produkcii, neoklasická teória rastu, kointegračná ekonometrická špecifikácia*

Abstract

The paper focuses on the econometric co-integration analysis of the firm's first order condition of labour. Using aggregate annual American data from the Jorgenson's sector tables we test a long-run factor share constancy hypothesis, which is an assumption for the neoclassical growth theory. Besides we do not reject the given hypothesis, we estimate a short-run elasticity of substitution. Its value is about 0.5.

Keywords: *factor shares, neoclassical growth theory, co-integration econometric estimation form*

1 ÚVOD

Caballero (1994) ponúkol ekonometrickú stratégiu odhadu dlhodobej elasticity substitúcie vstupov založenú na kointegračnej špecifikácii podmienky prvého rádu pre vstup kapitál firmy maximalizujúcej zisk. Avšak Chirinko a Mallick (2011) upozornili na problematickosť tejto stratégie. Autori uvažujú kointegračný vzťah v tvare

$$(y_t - x_t) = \gamma_0 + \gamma_1 (p_t^x - p_t^y) + v_t \quad (1)$$

kde y_t , x_t , p_t^y and p_t^x sú prirodzené logaritmy množstiev a cien produkcie a vstupu, parameter γ_1 je elasticita substitúcie vstupov podľa Caballera (1994) a v_t je náhodná zložka, pre ktorú sú splnené základné predpoklady lineárneho ekonometrického modelu. Neoklasická teória rastu predpokladá, že podiel vstupu x na produkcii y je $p_t^x + x_t - p_t^y - y_t$ je z dlhodobého hľadiska

¹ Príspevok vznikol s podporou projektu VEGA 1/0444/15 "Ekonometrická analýza produkčných možností ekonomiky a trhu práce na Slovensku" a tiež s podporou projektu VEGA 1/0285/14 „Regionálne modelovanie ekonomického rastu krajín EÚ s dôrazom na modely priestorovej ekonometrie“.

konštantný. Avšak, ako upozornili Chirinko a Mallick (2011), jeho dosadením do kointegračného vzťahu (1) je jeho konštantnosť splnená vtedy a len vtedy, keď sa parameter $\gamma_1 = 1$. Tento parameter nemôže byť interpretovaný ako dlhodobá elasticita vstupov.

Tri prípady zlučiteľné s teóriou uvádzajú Chirinko a Mallick (2011) pre možné interpretácie parametra γ_1 .

- Existuje kointegračný vzťah (1). Tento prípad sa môže vyskytnúť, ak uvažujeme vstup práca. Podľa neoklasickej teórie rastu je podiel práce na produkcii konštantný a relatívne ceny sú v dlhodobom vzťahu s relatívnymi množstvami. V takom prípade $\gamma_1 = 1$.
- Neexistuje kointegračný vzťah (1). Tento prípad sa môže vyskytnúť, ak uvažujeme vstup kapitál. Podľa neoklasickej teórie rastu je podiel kapitálu na produkcii konštantný, ale relatívne ceny sú dlhodobo nezávislé od relatívnych množstiev.
- Neexistuje kointegračný vzťah (1), avšak medzi premennými existuje iný kointegračný vzťah nezahrnutý v rovnici (1).

V práci sme sa zamerali na kointegračnú analýzu podmienky prvého rádu firmy pre prácu. Kointegračnou analýzou (1), využitím agregovaných amerických ročných Jorgensonových (2008) údajov cien a množstiev práce a produkcie v období 1960 – 2005, odhadneme parameter γ_1 blízky jednotke, z čoho vyplýva silná indícia, že podiel práce na produkcii je dlhodobo konštantný.

2 ÚDAJE

Pre odhad kointegračného vzťahu (1) analýzu sme využili Jorgensonové (2008) sektorové tabuľky. Množstvá vstupu práca a produkcie sme po sektoroch agregovali súčtami. Cenové indexy vstupu práca, resp. produkcie sme agregovali váženým priemerovaním s váhami množstiev vstupu práca, resp. produkcie.

Tabuľka 1. Korelogram relatívnych cien $p_t^l - p_t^y$, (Výstup z E-Views).

Sample: 1960 2005
Included observations: 46

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
. *****	. *****	1	0.912	0.912	40.847	0.000
. *****	. .	2	0.816	-0.097	74.277	0.000
. *****	. .	3	0.716	-0.074	100.61	0.000
. ****	. .	4	0.618	-0.047	120.67	0.000
. ****	. .	5	0.532	0.019	135.93	0.000
. ***	. .	6	0.455	-0.014	147.36	0.000
. ***	. .	7	0.385	-0.019	155.76	0.000
. **	. .	8	0.328	0.020	162.00	0.000
. **	. .	9	0.288	0.055	166.93	0.000
. **	. .	10	0.254	-0.005	170.89	0.000
. **	. .	11	0.230	0.023	174.23	0.000
. *	. .	12	0.210	-0.003	177.09	0.000
. *	. .	13	0.193	0.006	179.57	0.000
. *	. .	14	0.167	-0.068	181.49	0.000
. *	. .	15	0.130	-0.072	182.69	0.000

Zdroj: vlastné výpočty.

Testy stacionarity premenných sme realizovali rozšíreným Dickeyho-Fullerovým testom (ADF), Phillipsovým-Perronovým testom (PP) a korelogramom. Phillipsov-Perronov test sme

uprednostnili v prípade, ak vysoké množstvo autoregresných členov v testovacej špecifikácii riešilo autokoreláciu. Lukáčik a Lukáčiková (2008) vyzdvihli význam testovania stacionarity v časových radoch. Konštatujeme, že logaritmus relatívnych množstiev $y_t - l_t$ sú integrované rádu 1, kde symbolom l_t označujeme vstup práca. Logaritmus relatívnych cien $p_t^l - p_t^y$ je nestacionárny v rovnici bez konštanty a trendu podľa oboch ADF a PP testov. Pohľad na korelogramo v tab. 1 prezrádza, že hodnoty autokorelácie logaritmu relatívnych cien $p_t^l - p_t^y$ sa pomaly znižujú s rastom oneskorení a preto konštatujeme, že aj táto premenná je integrovaná rádu 1.

3 METODOLÓGIA

Pri testoch kointegrácie nevyužívame dvojstupňovú metódu Engla a Grangera (1987) pre nemožnosť priameho testovania parametrov dlhodobého vzťahu. Reziduál dlhodobej rovnováhy nechceme odhadovať samostatne. Pred Johansenovou (1995) procedúrou preferujeme dynamickú metódu najmenších štvorcov, ktorú navrhli Stock a Watson (1993) (pozri tiež prácu Lukáčikovej; 2013), lebo uvažujeme iba jeden kointegračný vzťah. Podľa postupu Stocka a Watsona môžeme špecifikáciu (1) zapísať v tvare:

$$\Delta(y_t - l_t) = \beta_0 + \beta_1 \Delta(p_t^l - p_t^y) + \beta_2 (y_{t-1} - l_{t-1}) + \beta_3 (p_{t-1}^l - p_{t-1}^y) + \beta_4 t + u_t \quad (2)$$

kde β_1 vyjadruje krátkodobú elasticitu substitúcie, kointegračný parameter je $-1 < \beta_2 < 0$. Parameter špecifikácie (1) γ_1 je: $\gamma_1 = -\beta_3/\beta_2$. Ak je tento parameter rôzny od jednej, z parametra β_4 vyjadruje možno odvodiť mieru rastu technológie rozširujúcu prácu (pozri Szomolányi, Lukáčik a Lukáčiková, 2015). Špecifikáciu (2) sme odhadli metódou najmenších štvorcov. Zaujímajú nás hodnoty parametrov β_1 , β_2 , β_3 , a β_4 . Ich odhady a odhady ich štandardných odchýlok sú v tab. 2. Ak je podiel práce na produkcii dlhodobo konštantný, platí: $\gamma_1 = 0$ a $\beta_4 = 0$. Združenú hypotézu testujeme využitím štatistiky podliehajúcej χ^2 štatistiky. Jej odhad je v poslednom riadku tab. 2.

Tabuľka 2. Odhad špecifikácie (2).

Parameter	Odhad	Št. odch.
β_1	0,513	0,092
β_2	-0,115	0,059
β_3	0,148	0,058
β_4	-0,001	0,001
$\chi^2 = 1,389$		

Zdroj: vlastné výpočty.

Parametre sú štatisticky významné na 5 % hladine významnosti, až na kointegračný parameter β_2 , ktorý je štatisticky významný na 10 % hladine významnosti a parameter trendu β_4 , z čoho, popri hodnote $\gamma_1 = -\beta_3/\beta_2$ blízkej 1, možno predvídať dlhodobú konštantnosť podielu práce na produkcii. V skutočnosti hodnota štatistiky χ^2 testu hypotézy dlhodobej konštantnosti podielu práce na produkcii je nízka ako kritická hodnota štatistiky. Hypotézu nezamietame. Ak zohľadníme reštrikcie vyplývajúce z hypotézy, $-\beta_3/\beta_2 = 1$ a $\beta_4 = 0$ a znova odhadneme prislúchajúcu ohraničenú špecifikáciu (2) získame odhady parametrov, ktoré sú štatisticky významné na 5 % hladine významnosti. V tejto špecifikácii je hodnota odhadu krátkodobej elasticity substitúcie 0,499.

Stacionaritu reziduálov sme testovali využitím ADF a PP testov postupom opísaným časti Údaje. Autokoreláciu sme testovali Breshovým-Godfreyovým testom. Konštatujeme, že reziduály odhadu sú stacionárne a neautokorelované.

4 ZÁVER

Na amerických údajov z Jorgensonových (2008) sektorových tabuliek sme potvrdili dlhodobú konštantnosť elasticity substitúcie vstupov, ktorá je predpokladom neoklasickej teórie rastu. Okrem toho sme odhadli krátkodobú elasticitu substitúcie vstupov. Jej hodnota sa pohybuje okolo 0,5.

Použitá literatúra

1. CABALLERO. R. J. 1994. Small Sample Bias and Adjustment Costs. *The Review of Economics and Statistics*. 76(1).
2. CHIRINKO. R. S. a MALLICK. D. 2011. Cointegration, Factor Shares, and Production Function Parameters. *Economics Letters*, 112(2), 205-206.
3. ENGLE, R. F. a Granger, C. W. J. 1987. Co-integration and error correction representation, estimation and testing. *Econometrica*, 55(2), 251-276.
4. LUKÁČIK, M. 2013. Parametre produkčnej funkcie ekonomiky. In: Zborník medzinárodného vedeckého seminára Nové trendy v ekonometrii a operačným výzkumu Bratislava: Ekonóm.
5. LUKÁČIK. M. a LUKÁČIKOVÁ. A. 2008. Význam testovania stacionarity v ekonometrii. *Ekonomika a informatika*. 6(1). 146-157.
6. LUKÁČIKOVÁ, A. 2013: Englova-Grangerova procedúra a dynamická metóda najmenších štvorcov. In: Zborník 15. medzinárodnej vedeckej konferencie AIESA – budovanie spoločnosti založenej na vedomostiach, Bratislava: Ekonóm.
7. SZOMOLÁNYI. K., LUKÁČIK. M. a LUKÁČIKOVÁ. A. 2015. Long-Run Elasticity of Substitution. In: Proceedings of 33rd international conference Mathematical methods in economics (s. 777-781), Plzeň: Faculty of Economics. University of West Bohemia.
8. JOHANSEN, S. 1995. Identifying Restrictions of Linear Equations with Applications to Simultaneous Equations and Cointegration. *Journal of Econometrics*. 69(1), 111-132.
9. JORGENSON. D. W. 2008. *35 Sector KLEM*. Harvard Dataverse.
10. STOCK, J. H. a WATSON, M. W. 1993. A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica* 61, s. 783-820.

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: karol.szomolanyi@euba.sk

ANALÝZA S&P 500 POMOCÍ MODELU STOCHASTICKÉ VOLATILITY

ANALYSIS OF S&P 500 USING STOCHASTIC VOLATILITY MODEL

Petra Tomanová

Abstrakt

Tato práce se zabývá odhadem volatility výnosů akcií S&P 500 pro roky 2005 až 2014. K tomuto účelu je použit model stochastické volatility využívající k odhadu parametrů Kalmanův filtr. Poté je odhad míry volatility modelován pomocí metody Kalman smoother.

Klíčová slova: Stochastická volatilita, S&P 500, Kalmanův filtr

Abstract

The aim of this paper is to estimate volatility of returns of S&P 500. The sample period spans from 2005 to 2014. The approach is based on stochastic volatility model. Kalman filter is utilized for the coefficients estimation. The estimates of volatility measure is modeled using Kalman smoother method.

Keywords: Stochastic Volatility, S&P 500, Kalman Filter

1 ÚVOD

Modely stochastické volatility se zejména používají v oblasti finanční matematiky pro ohodnocování derivátů. Jejich základní charakteristikou je fakt, že volatilita je modelována jako nepozorovaná latentní proměnná. Srovnání modelů stochastické volatility lze například nalézt v [2]. V této práci je aplikován model stochastické volatility využívající metody Kalmanův filtr a Kalman smoother na výnosy akcií S&P 500.

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 μ je průměrný výnos, $\sigma(t)$ 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 ξ je konstanta, $0 < \lambda < 1$, σ_η je volatilita volatility (striktně pozitivní) a $B(t)$ je standardizovaný Brownův 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 ε_t a η_t .

2.2 Kalmanův filter

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 $\alpha_1 \sim N(a_1, P_1)$. ζ_{1t} a ζ_{2s} jsou nezávislé pro všechny t, s a nezávislé od α_1 . State vektor α_t je nepozorovaný a matice T_t, Z_t, R_t, Q_t, H_t jsou fixní v čase t .

Nepozorovaný vektor α_t lze "odhadnout" z pozorování za použití Kalmanova filtru [1]:

$$\begin{aligned} v_t &= y_t - Z_t a_t, \\ F_t &= Z_t P_t Z_t' + H_t, \\ K_t &= T_t P_t Z_t' F_t^{-1}, \\ a_{t+1} &= T_t a_t + K_t v_t, \\ P_{t+1} &= T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t' \end{aligned}$$

pro $t = 1, \dots, n$ a počátečními hodnotami a_1 a P_1 . Platí $Y_t = \{y_1, \dots, y_t\}$,

$$a_{t+1} = \mathbb{E}(\alpha_{t+1} | Y_t), \quad P_{t+1} = \mathbb{V}\text{ar}(\alpha_{t+1} | Y_t).$$

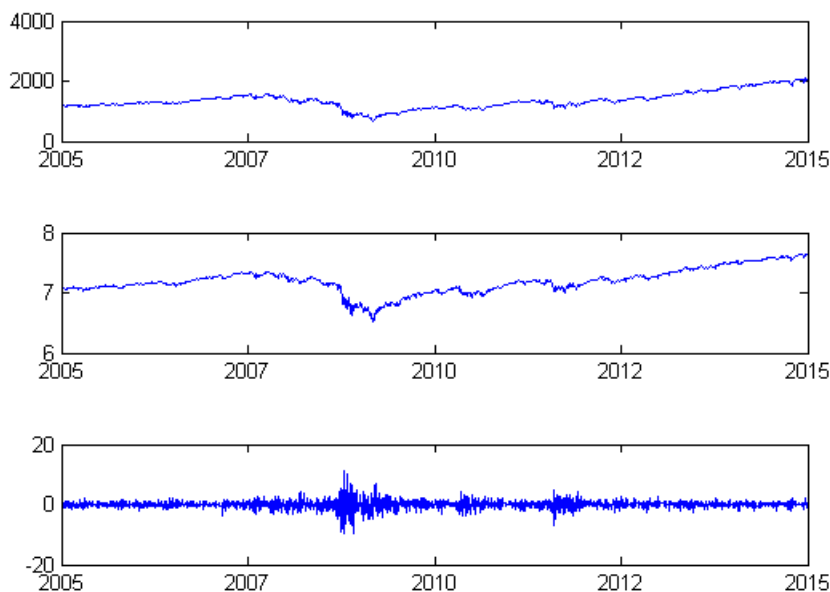
Poté co jsou filtrované odhady spočítány, lze použít následující rekurzi (Kalman smoother) začínající posledním pozorováním a končící prvním [1]:

$$\begin{aligned} r_{t-1} &= F_t^{-1} v_t + L_t r_t, \\ N_{t-1} &= F_t^{-1} + L_t' N_t L_t, \\ \hat{\alpha}_t &= a_t + P_t r_{t-1}, \\ V_t &= P_t - P_t N_{t-1} P_t. \end{aligned}$$

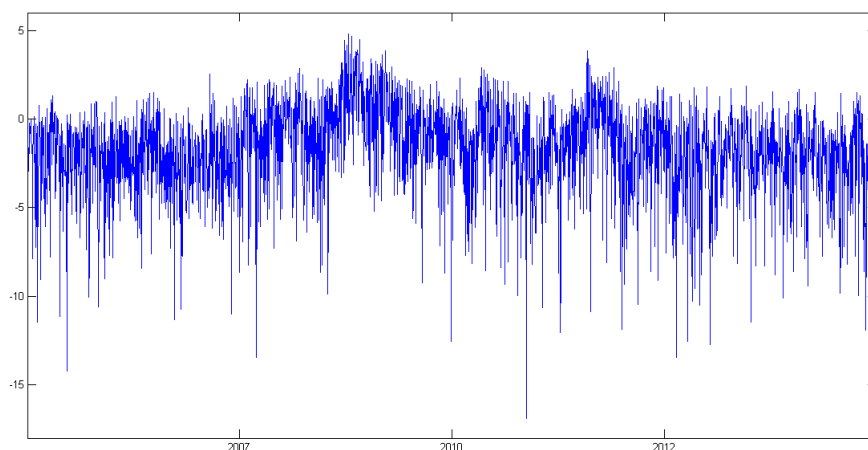
3 ANALÝZA S&P 500

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 staženy z www.finance.yahoo.com. Pro účely analýzy byly z cen akcií spočteny procentuální logaritmické výnosy. Původní a transformované časové řady na logaritmické výnosy jsou zobrazeny v grafu 1 a transformovaná časová řada pomocí substituce za účelem linearizace modelu je zobrazena v grafu 2. Tabulka 1 zobrazuje popisné statistiky časových řad.



Obrázek 1: Graf (i) cen, (ii) logaritmických cen (iii) logaritmických výnosů of S&P 500.



Obrázek 2: Transformovaná časová řada $x_t = \log(y_t - \mu)^2$.

3.2 Výsledky

Pro odhady koeficientů je předpokládáno, že $u_t = \log \varepsilon_t^2 \sim \text{NID}(0, \sigma_u^2)$ a $\eta_t \sim \text{NID}(0, \sigma_\eta^2)$, takže je možné odhadnout neznámé koeficienty σ_u^2 , σ_η^2 , ω a ϕ metodou Quasi Maximum Likelihood,

	Mean	Median	Min	Max	SD	Skewness	Ex. Kurtosis
Price	1352	1313	676,5	2090,6	278,3	0,5485	0,2318
Log price	7,1887	7,1801	6,5170	7,6452	0,2053	-0,1003	0,2931
Log returns	0,0000	0,0579	-9,4909	10,9358	1,2887	-0,3340	11,0144

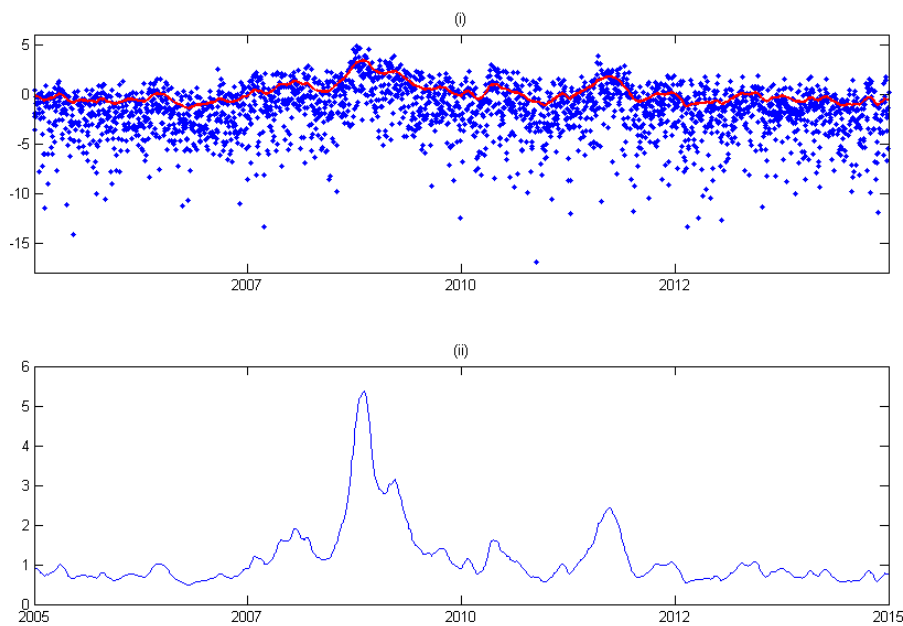
Tabulka 1: Popisné statistiky cen, logaritmických cen a logaritmických výnosů S&P 500.

kde log-likelihood je získána z Kalmonového filtru. Model je implementován v Matlab 2014a. Omezení parametrů bylo stanoveno následovně: $\sigma_u^2 > 0$, $\sigma_\eta^2 > 0$, $\omega > 0$ a $0 < \phi < 1$.

Hodnoty odhadnutých koeficientů:

$$\sigma_u^2 = 5.7399, \quad \sigma_\eta^2 = 0.0180, \quad \omega = 1.0011e - 07, \quad \phi = 0.9977.$$

Získané odhadnuté koeficienty jsou použity pro výpočet vyhlazeného průměru h_t na základě aproximativního modelu pro x_t za použití metody Kalman smoother. Výsledky jsou zobrazeny v grafu 3.



Obrázek 3: (i) logaritmická řada y_t^2 (modře, tečkovaně) s vyrovnanými odhady h_t (červená čára); (ii) vyrovnané odhady míry volatility $\exp(h_t/2)$.

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 Kalmanův filtr a Kalman smoother. Další práce se bude zabývat srovnáním uvedené metody s jejími modifikacemi: model stochastické volatility za použití importance sampling a simulation smoothing algorithm, modely stochastické volatility předpokládající t -rozdělení náhodných složek, stochastický model pro odhad vyhlazeného modusu h_t .

Poděkování

Tento článek byl napsán v rámci grantu IGS F4/63/2016 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. 5 Stochastic volatility. *Handbook of statistics*, 1996, 14: 119-191.