ISSN 2029-7564 (online)
SOCIALINĖS TECHNOLOGIJOS
SOCIAL TECHNOLOGIES
2011, 1(2), p. 301–315.

NEKILNOJAMOJO TURTO VERTĖS MODELIAVIMAS TAIKANT DICHOTOMINIO TESTAVIMO METODIKĄ

Aleksandras Krylovas

Mykolo Romerio universitetas, Lietuva, krylovas@mruni.eu

Natalja Kosareva

Vilniaus Gedimino technikos universitetas, Lietuva, natalja.kosareva@vgtu.lt

Laura Gudelytė

Mykolo Romerio universitetas, Lietuva, l.gudelyte@mruni.eu

Tadas Laukevičius

Mykolo Romerio universitetas, Lietuva, tadas@mruni.eu

Abstraktas

Tikslas. *Autoriai siekia sukurti algoritmą, kuriuo automatiškai atliekami skaičiavimai ir klasterizuojami Vilniaus miesto mikrorajonai pagal normuotą kainą atsižvelgiant į turimą statistinę informaciją.*

Metodologija. *Klasterizavimo metodas pasitelkiant pačių sukurtą algoritmą.*

Rezultatai. *Autoriai pristato sukurtą algoritmą, kuriuo automatiškai parenkami informatyviausi dichotominiai diagnostiniai operatoriai. Šie operatoriai yra optimalūs, kadangi jų pagrindu atlikta mikrorajonų klasterizacija yra artimiausia mikrorajonų klasterizacijai pagal normuotą kainą.*

Praktinė reikšmė. *Kuriamas determinuotas vertės rangavimo modelis.*

ISSN 2029-7564 (online)
https://www.mruni.eu/lt/mokslo_darbai/st/apie_leidini/
https://www.mruni.eu/en/mokslo_darbai/st/apie_leidini/index.php

***Originalumas**. Modeliavimo metodika nėra originali, tačiau ji taikoma tokioje srityje, kurioje išplėtoti statistiniai, o ne determinuoti matematiniai modeliai.*

***Tyrimo tipas**: tyrimo pristatymas.*

***Raktažodžiai**: matematinis modeliavimas, indikatorių konstravimas, dichotominis testas.*

1. Įvadas

Nekilnojamojo turto vertinimas yra svarbus jo savininkams, pirkėjams, bankams, lizingo įmonėms, investuotojams – perkant būstą, sudarant paskolos sutartį, ieškant investavimo galimybių, rengiant finansinę atskaitomybę, nustatant nuomos kainą ar kilus civiliniam ginčui. Teisingas nekilnojamojo turto vertės nustatymas yra svarbus ir turto pirkėjams, ir pardavėjams, bankams, suteikiantiems paskolas, nekilnojamąjį turto registrą tvarkančioms įstaigoms, draudimo įstaigoms. Pagal 1999 m. gegužės 25 d. LR turto ir verslo vertinimo pagrindų įstatymą Nr. VIII-1202 turto kaina valstybės įstaigose nustatoma lyginamosios vertės (pardavimo vertės analogų) metodu (toliau tekste – lyginamosios vertės metodas), kurio esmė yra palyginimas, t. y. turto rinkos vertė nustatoma palyginus analogiškų objektų faktinių sandorių kainas, kartu atsižvelgiant į nedidelius vertinamo turto bei jo analogo skirtumus. Panašiai nekilnojamasis turtas vertinamas ir privataus sektoriaus įmonėse, kur įprastai taikomi įvairiausi papildomi metodai, kurie dažniausiai nėra automatizuoti. Kita vertus, atsiranda vis daugiau automatizuotų metodų, kuriais galima vertinti įvairių rūšių nekilnojamąjį turtą, tačiau jų tikslumas itin priklauso nuo išorinių sąlygų ir neapibrėžtumo rinkoje, kurį sudėtinga kiekybiškai išmatuoti. Užsienio šalyse paplitę ekonometriniai komercinio ir nekomercinio nekilnojamojo turto vertinimo modeliai, pagrįsti panelinių duomenų analizės metodais (An X, Deng Y, Fisher J.D., 2011; Chan, N., 2001; Scott R. Muldavin, 2010). Kitas vertinimo būdas – daugiakriteriniai (MCDM) ryškiųjų ir neryškiųjų aibių metodai (Antuchevičienė, J., Zakarevičius, A., Zavadskas, 2011). Šiame straipsnyje siekiant apytikriai nustatyti gyvenamojo būsto vieno kvadratinio metro kainą atitinkamame miesto mikrorajone, surūšiuojant mikrorajonus pagal kainą, taikomas deterministinis, t. y. dichotominio testo vertinimo, modelis. Tuo tikslu ankstesniuose autorių darbuose (Krylovas A., Kosareva N., 2010; Krylovas A., Kosareva N., 2009; Krylovas A., Kosareva N., 2008) buvo paaiškinta, kaip turint apriorinę informaciją apie tam tikrų testo klausimų funkcijas ir apie tiriamojo populiacijos požymio skirstinį, sukonstruoti testą (indikatorių), maksimizuojantį gaunamą kiekį informacijos. Tačiau autorių darbe (Krylovas A., Kosareva N., Gudelytė, L., 2011) buvo išdėstyta tik bendra šios metodikos pritaikymo socialiniams indikatoriams konstruoti schema. Šiame darbe siūlomas indikatorių konstravimo algoritmas, kuris taikomas analizuojant atskirą modeliavimo pavyzdį: Vilniaus miesto mikrorajonų pasiskirstymą pagal vidutinę būsto kainą. Modelyje daroma prielaida, kad indikatorius matuoja tam tikrą latentinį kintamąjį $p \in [0;1]$, o klausimų charakteristinės funkcijos[1] $k(p):[0,1]\to[0,1]$ yra nemažėjančios (Krylovas A., 2009). Tačiau socialiniuose reiškiniuose kintamasis p dažnai būna tik tam tikro susitarimo apie objektų ranginę tvarką rezultatas, todėl ši prielaida gali negalioti.

2. Nekilnojamojo turto kainos modeliavimo pavyzdys

Šiame skyriuje parodoma, kaip nustatoma apytikrė butų, esančių skirtinguose Vilniaus mikrorajonuose, kaina remiantis dichotominio testo vertinimo modeliu. Kadangi rajonų kokybinės savybės matuojamos skirtingų dydžių skalėse, klasterių sudarymui rajonų kokybinių savybių vektoriaus $(x_1, x_2,..., x_m)$ komponenčių reikšmės normuojamos taikant formulę

$$x_j = \frac{x_j^{fakt} - x_j^{\min}}{x_j^{\max} - x_j^{\min}}; \qquad (1)$$

čia $x_j^{\min} - j$ –tojo rodiklio minimali reikšmė, $x_j^{\max} - j$ – tojo rodiklio maksimali reikšmė, x_j^{fakt} – stebima *j*-tojo rodiklio nenormuota reikšmė, $j = 1,2,...., m$. Tuomet baigtinė mikrorajonų aibė $A = \{a_1, a_2,..., a_n\}$ suskaidoma į nesikertančius blokus (klasterius) (Krylovas A., 2009).

$$K_{0.}, K_1,..., K_m \subset A, K_i \neq \varnothing, K_i \cap K_j = \varnothing, \bigcup_{j=0}^{m} K_j = A.$$

Klasteriai K_j apibrėžia aibėje A dalinę tvarką pagal gyvenamojo ploto vieno kvadratinio metro kainą (Krylovas A., 2009): $x < y \Leftrightarrow x \in K_i, y \in K_j, i < j$.

Nagrinėkime 1 pav. pateiktus Vilniaus miesto mikrorajonus, sunumeruotus didėjančia tvarka pagal vidutinę ketverių metų (2007–2010 m.) kvadratinio metro būsto kainą, ir suskirstykime juos į keturis klasterius:

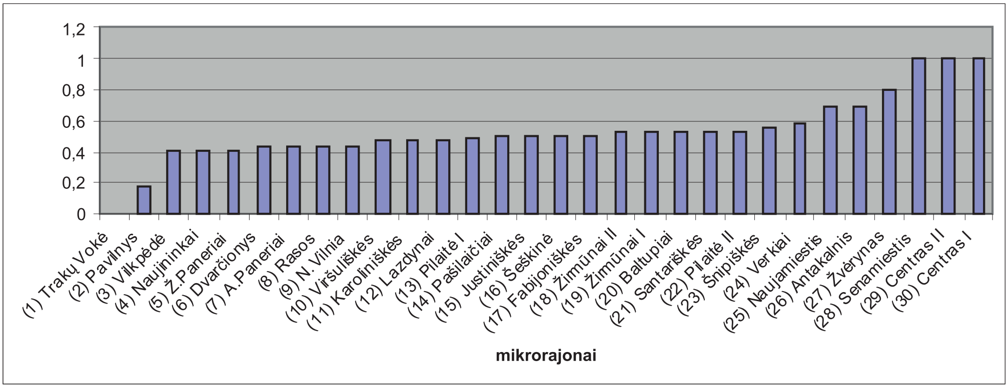
K_0= {Trakų Vokė (1), Pavilnys (2), Ž. Paneriai (3), Naujininkai (4), Vilkpėdė (5), Dvarčionys (6), A. Paneriai (7), Rasos (8)};

K_1 = {N. Vilnia (9), Viršuliškės (10), Lazdynai (11), Karoliniškės (12), Pilaitė I (13), Šeškinė (14), Justiniškės (15)};

K_2 = {Pašilaičiai (16), Fabijoniškės (17), Žirmūnai I (18), Žirmūnai II(19), Baltupiai (20), Santariškės (21), Pilaitė II (22)};

K_3 = {Šnipiškės (23), Verkiai (24), Naujamiestis (25), Antakalnis (26), Žvėrynas (27), Centras I (28), Centras II (29), Senamiestis (30)}.

1 T. y. tikimybės teisingai atsakyti į atitinkamą testo klausimą.



1 pav. Mikrorajonų suskirstymas pagal vidutinę butų kainą

Apibrėžkime diagnostinį operatorių: kiekvienam aibės A elementui a, priklausomai nuo testo (indikatoriaus) klausimo, priskirkime reikšmę $p_\alpha(a;x)$ – nulį arba vienetą. Funkcija p aibę A suskaido į du klasterius:

$$A^0 = \{a \in A : p_\alpha(a;x) = 0\};$$
$$A^1 = \{a \in A : p_\alpha(a;x) = 1\}$$

čia

$$p_\alpha(a;x) = \begin{cases} 1, & \text{jei } x \geq \alpha \\ 0, & \text{jei } x < \alpha \end{cases}, \tag{2}$$

jei tarp rajono kainos ir faktoriaus x yra teigiama koreliacija,

$$p_\alpha(a;x) = \begin{cases} 1, & \text{jei } x \leq \alpha \\ 0, & \text{jei } x > \alpha \end{cases}, \tag{3}$$

jei tarp rajono kainos ir faktoriaus x yra neigiama koreliacija.

Pažymėkime C_i klasterį, sudarytą iš pirmųjų $i+1$ klasterių $K_i: C_i = K_0 \cup K_1 \cup \ldots \cup K_i i = 0,1,2,\ldots,m-1$. Tuomet funkcija $p_\alpha(a;x): A \to \{0;1\}$ yra dichotominis diagnostinis operatorius, kai $\forall i \in 0,1,2,\ldots,m-1$ teisinga nelygybė

$$\left|A^0 \cap C_i\right| + \left|A^1 \cap (A \setminus C_i)\right| \geq \left[\frac{n+1}{2}\right]. \tag{4}$$

(4) diagnostinio operatoriaus teisingai priskirtų reikšmių 0 ir 1 skaičius turi būti ne mažesnis nei pusė visų priskirtų reikšmių. Kitaip tariant, elementų $x_i \in K_i, x_j \in K_j, i < j$, kuriems galioja nelygybė $p_\alpha(a;x_i) \leq p_\alpha(a;x_j)$ yra ne mažiau negu elementų, kuriems ši nelygybė negalioja. Tai yra diskretusis analogas tikimybinių reikalavimų (Krylovas A., Kosareva N., 2010; Krylovas A., Kosareva N., 2009; Krylovas A., Kosareva N.,

2008) diagnostiniams operatoriams, kuris turi validumo prasmę: sudarytas iš diagnostinių operatorių $p_\alpha(a; x)$ testas (indikatorius) turi matuoti tai, kam jis sukurtas – apibrėžti naują mikrorajonų klasterizaciją pagal kainą $\widetilde{K}$, kiek įmanoma artimesnę klasterizacijai K (Krylovas A., Kosareva N., Gudelytė, L., 2011).

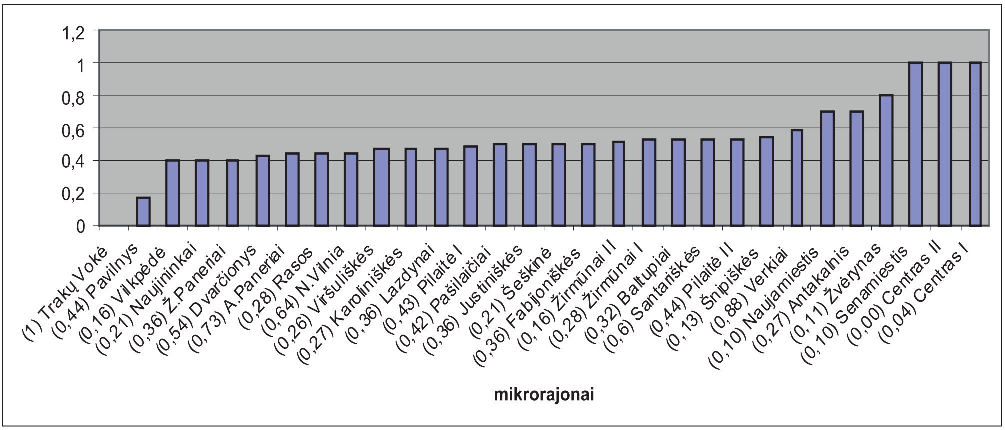
Analizuokime diagnostinio operatoriaus pavyzdį, kai x – mikrorajono atstumo nuo miesto centro standartizuota reikšmė (2 pav.). Diagnostinis operatorius turi atitikti (4) nelygybę ir formulę (3), nes tarp rajono kainos ir x yra atvirkštinė priklausomybė – rajonai, kurie yra toliau nuo miesto centro, yra pigesni.

K_0 = {Trakų Vokė (1), Pavilnys (0,44), Vilkpėdė (0,16), Naujininkai (0,21), Ž. Paneriai (0,36), Dvarčionys (0,54), A. Paneriai (0,73), Rasos (0,28)};

K_1 = {N. Vilnia (0,64), Viršuliškės (0,26), Karoliniškės (0,27), Lazdynai (0,36), Pilaitė I (0,43), Pašilaičiai (0,42), Justiniškės (0,36)};

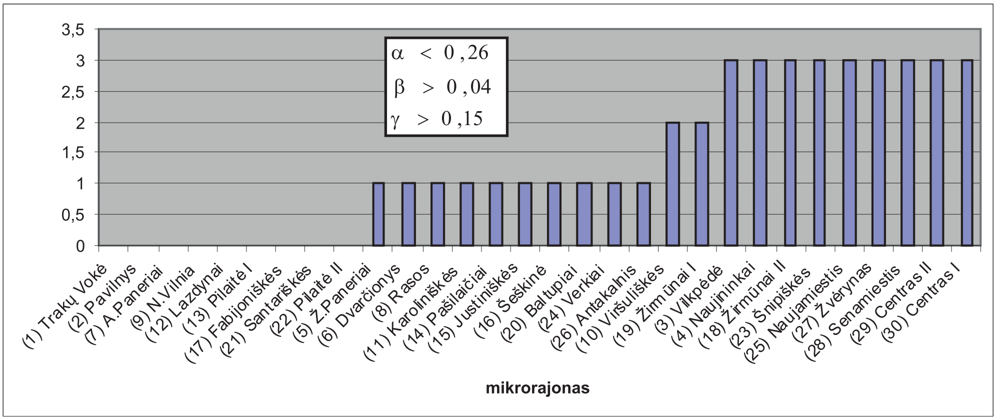
K_2 = {Šeškinė (0,21), Fabijoniškės (0,36), Žirmūnai II (0,16), Žirmūnai I (0,28), Baltupiai (0,32), Santariškės (0,6), Pilaitė II (0,44)};

K_3 ={Šnipiškes (0,13), Verkiai (0,88), Naujamiestis (0,10), Antakalnis (0,27), Žvėrynas (0,11), Senamiestis (0,10), Centras II (0,00), Centras I (0,04)}.



2 pav. Mikrorajonų suskirstymas pagal vidutinę kainą, nurodant jų standartizuotą atstumą nuo miesto centro

Pažymėkime $|A|$ – aibės A elementų skaičių. Iš klasterio $C_0 = K_0$ atrenkami nariai, kurių atstumo nuo centro reikšmė yra didesnė nei 0,26. Gausime $\left|A^0 \cap C_0\right| = 6$. Iš likusių klasterių K_1, K_2, K_3 suskaičiuojame, kiek yra mikrorajonų, kurių atstumas nuo centro yra mažesnis nei 0,26 $\left|A^1 \cap (A \setminus C_0)\right| = 9$. Šiuo atveju, kai $i = 0$, visai klasterizacijai pritaikę (4) formulę, gauname: $6+9=\left\lceil \frac{30+1}{2} \right\rceil$. Matome, kad pasirinkta funkcija $1_{\{x\leq 0.26\}}$ atitinka (4) nelygybę. Panašiai tikriname (4) nelygybę klasteriams C_1 ir C_2.



3 pav. Mikrorajonų suskirstymas pagal atstumą nuo miesto centro, darbo vietų tankį ir gatvių tankį

3. Klasterizacijos nesuderinamumo laipsnis

Diagnostinių operatorių rinkinio atveju, kai turime tris diagnostinius operatorius $p_\alpha(a;x)$; $p_\beta(a;y)$; $p_\gamma(a;z)$, čia x – mikrorajono atstumo nuo miesto centro, y – darbo vietų tankio, z – gatvių tankio standartizuotos reikšmės, sudarome mikrorajonų (jie sunumeruoti) klasterizaciją pagal tris kintamuosius (3 pav.). Nauja klasterizacija sudaryta pagal trijų diagnostinių operatorių sumos reikšmę (nuo 0 iki 3):

$\tilde{K}_0 =$ {Pilaitė II, Pavilnys, A. Paneriai, Trakų Vokė, Lazdynai, Pilaitė I, N. Vilnia, Fabijoniškės, Santariškės};

$\tilde{K}_1 =$ {Ž. Paneriai, Rasos, Karoliniškės, Šeškinė, Justiniškės, Antakalnis, Pašilaičiai, Baltupiai, Dvarčionys, Verkiai};

$\tilde{K}_2 =$ {Viršuliškes, Žirmūnai II};

$\tilde{K}_3 =$ {Centras I, Centras II, Naujininkai, Naujamiestis, Vilkpėdė, Senamiestis, Žvėrynas I, Žvėrynas II, Šnipiškės}.

Pagal aprašytą metodiką [4] sudaromi klasterizacijų K ir $\widetilde{K}$ neatitikimai: $D_j = \{a \in A : a \in K_i \cap \widetilde{K}_l,\ |i-l| = j\}$, kai $j = 1,\ldots,m-1$; čia $A = \{1,2,\ldots,30\}$ – mikrorajonų aibė. Idealiuoju atveju indikatorius $I(A) = \widetilde{K}$ apibrėžia identiškus suskirstymus (t. y. tą pačią klasterizaciją) pagal tris testo klausimus ($\widetilde{K}$) ir kainą (K), todėl šiuo atveju $D_j = \{a \in A : a \in K_i \cap \widetilde{K}_l,\ |i-l| = j\}$. Dydis $|D_0|$ parodo, kiek yra mikrorajonų, patenkančių į klasterius su tais pačiais numeriais, $|D_1|$ – kiek patenka į klasterius, kurių numeriai skiriasi per vieną, $|D_2|$ – kiek mikrorajonų patenka į klasterius, kurių numeriai skiriasi per du ir t. t. Klasterizacijos nesuderinamumo laipsnį išmatuojame funkcija (5):

$$S(I(A),K) = \sum_{j=1}^{m-1} j \cdot |D_j|. \tag{5}$$

Tuomet naudodamiesi [6] darbe pasiūlytu kriterijumi gauname tokį klasterizacijų $\widetilde{K}$ ir K neatitikimo laipsnį: $S(I(A),K) = 1 \cdot 10 + 2 \cdot 5 + 3 \cdot 2 = 26$.

4. Automatinės klasterizacijos algoritmas

Tiksliausią indikatorių $I_0(A)$ gausime, kai

$$S(I(A),K) = \min_{\{p_1,p_2,\ldots,p_m\}} S(\{p_1,\ldots,p_m\},\ K). \tag{6}$$

Idealiuoju atveju $S(I(A),K) = 0$ ir $\dfrac{|D_0|}{|A|} = 1$.

Siekiant nustatyti geriausią klasterizaciją, kuriamas skaičiavimo algoritmas.

Nagrinėkime duomenų matricą A.

$$A = \begin{pmatrix} k_1 & x_1 & y_1 \\ k_2 & x_2 & y_2 \\ \ldots & \ldots & \ldots \\ k_n & x_n & y_n \end{pmatrix}, \tag{7}$$

kurioje $k_j \in \{0,1,\ldots,k-1\}$ – klasterio numeris, $x_j,\ y_j,\ 0 \le x_j \le 1,\ 0 \le y_j \le 1$ – standartizuotos požymių x ir y reikšmės. Tuomet konstruokime diagnostinį operatorių

$$\mathrm{x}_{\alpha}(a;x)=\begin{cases}1, & \mathrm{x}\geq\alpha,\\ 0, & \mathrm{x}<\alpha,\end{cases} \tag{8}$$

Pastaba 1. *Kai tarp klasterių numeracijos k ir požymio x yra neigiama koreliacija, atliekama kintamojo x reikšmių* x_j *transformacija* $\tilde{x}_j = 1 - x_j$.

$$A_j = \begin{pmatrix} t_1 & \tilde{k}_1 & \tilde{x}_1 \\ t_2 & \tilde{k}_2 & \tilde{x}_2 \\ \ldots & \ldots & \ldots \\ t_n & \tilde{k}_n & \tilde{x}_n \end{pmatrix}; \tag{9}$$

čia t_j – elementų numeriai, o $\tilde{x}_j$ – x_j stulpelis, surūšiuotas nemažėjimo tvarka. Tikriname, ar pasirinkta α reikšmė yra tinkama, t. y. ar kriterijus $p_{\alpha}(a;x)$ atitinka monotoniškumo principą. Skaičiuojame, kiek skirtingų klasterių elementų yra virš α ir kiek po α.

$$T_{\alpha}^{0} = \sum_{\tilde{x}_j \leq \alpha;\, \tilde{k}_j = 0} 1 \tag{10}$$

– elementų, priklausančių klasteriui C_0, skaičius viršutinėje matricos A_j dalyje.

$$T_{\alpha}^{1,2,\ldots} = \sum_{\tilde{x}_j > \alpha;\, \tilde{k}_j \neq 0} 1 \tag{11}$$

– elementų, priklausančių klasteriui $\{A \setminus C_0\}$, skaičius apatinėje matricos A_j dalyje. Tikriname, ar teisinga nelygybė

$$T_{\alpha}^{0} + T_{\alpha}^{1,2,\ldots} \geq \left[\frac{n+1}{2}\right] \tag{S_0}$$

$$T_{\alpha}^{0,1} = \sum_{\tilde{x}_j \leq \alpha;\, \tilde{k}_j = 0 \vee \tilde{k}_j = 1} 1 \tag{12}$$

– elementų, priklausančių klasteriui C_1, skaičius viršutinėje matricos A_j dalyje.

$$T_{\alpha}^{2,3,\ldots} = \sum_{\tilde{x}_j > \alpha;\, \tilde{k}_j \neq 0 \,\&\, \tilde{k}_j \neq 1} 1 \tag{13}$$

– elementų, priklausančių klasteriui $\{A \setminus C_0\}$, skaičius apatinėje matricos A_j dalyje. Tikriname, ar teisinga nelygybė

$$T_\alpha^{0,1} + T_\alpha^{2,3,...} \geq \left[\frac{n+1}{2}\right] \qquad (S_{0,1})$$

...

$$T_\alpha^{0,1...l} = \sum_{\tilde{x}_j \leq \alpha; \tilde{k}_j = 0 \vee \tilde{k}_j = 1 \vee ... \vee \tilde{k}_j = l} 1 \tag{14}$$

$$T_\alpha^{l+1;l+2,....} = \sum_{\tilde{x}_j > \alpha; \tilde{k}_j \neq l+1 \& \tilde{k}_j \neq l+2 \& ... \& \tilde{k}_j \neq k-1} 1 \tag{15}$$

$$T_\alpha^{0,1,...;l} + T_\alpha^{l+1;l+2,....} \geq \left[\frac{n+1}{2}\right] \qquad (S_{0,1,...,l})$$

...

Sąlygos $S_0, S_{0,1},..., S_{0,1,...,k-1}$ tikrinamos su visomis skirtingomis

$$\alpha_i = \frac{\tilde{x}_i + \tilde{x}_{i+1}}{2}, \; i = 1,2,..., n-1 \cdot \tag{16}$$

Panagrinėkime pavyzdį, kuriame paimta 10 mikrorajonų $(n = 10)$. Pirmame stulpelyje turime suskirstymą į tris klasterius pagal kainą, antrame ir trečiame stulpeliuose yra standartizuotos konkrečių faktorių, apibūdinančių mikrojajonus, reikšmės, atitinkamai atstumas nuo miesto centro ir gatvių tankis.

$$A^T = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 2 & 1 \\ 0{,}1 & 0{,}6 & 0{,}9 & 0{,}4 & 0{,}5 & 0{,}6 & 0{,}2 & 0{,}3 & 0{,}8 & 0{,}7 \\ 0{,}3 & 0{,}7 & 0{,}8 & 0{,}4 & 0{,}6 & 0{,}3 & 0{,}3 & 0{,}2 & 0{,}7 & 0{,}7 \end{pmatrix}.$$

Pasirinkime visas α reikšmes pagal (16) formulę. Jos gali būti tokios:

$$\alpha_j \in \{0{,}15; 0{,}25; 0{,}35; 0{,}45; 0{,}55; 0{,}65; 0{,}75; 0{,}85\} \cdot$$

Tuomet pasirinkime $\alpha_1 = 0{,}15$. Matricos A trečiojo stulpelio elementus surikiuokime didėjimo tvarka. Antrame stulpelyje yra sena klasterizacija, pirmame – elemento numeris. Reikia patikrinti, ar $\alpha_1 = 0{,}15$ atitinka (4) formulę.

$$A_1^T = \begin{pmatrix} 1 & 7 & 8 & 4 & 5 & 2 & 6 & 10 & 9 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 2 \\ 0,1 & 0,2 & 0,3 & 0,4 & 0,5 & 0,6 & 0,6 & 0,7 & 0,8 & 0,9 \end{pmatrix}$$

1. $T_{0,15}^{0} = \{1\}$,
   $T_{0,15}^{1,2} = \{2,3,4,5,6,9,10\}$,

   $$\left|\mathrm{T}_{0,15}^{0}\right| + \left|\mathrm{T}_{0,15}^{1,2}\right| = 1 + 7 = 8 \geq \left[\frac{n+1}{2}\right] = 5.$$

2. $T_{0,15}^{0,1} = \{1\}$,
   $T_{0,15}^{2} = \{3,9\}$,
   $\left|T_{0,15}^{0,1}\right| + \left|T_{0,15}^{2}\right| = 1 + 2 = 3 < 5.$

Matome, kad $\alpha_1 = 0,15$ – netinkama reikšmė.

Imame kitą $\alpha_3 = 0,35$ reikšmę ir vėl atliekame analogiškus skaičiavimus, kaip ir prieš tai.

1. $T_{0,35}^{0} = \{1,7,8\}$,
   $T_{0,35}^{1,2} = \{2,3,4,5,6,9,10\}$,
   $\left|T_{0,35}^{0}\right| + \left|T_{0,35}^{1,2}\right| = 3 + 7 = 10 \geq 5.$

2. $T_{0,35}^{0,1} = \{1,7,8\}$,
   $T_{0,35}^{2} = \{3,9\}$,
   $\left|T_{0,35}^{0,1}\right| + \left|T_{0,35}^{2}\right| = 3 + 2 = 5 \geq 5.$

Šiuo atveju gauname, kad $\alpha_3 = 0,35$ – tinkama reikšmė.

Patikrinkime dar vieną reikšmę $\alpha_6 = 0,65$:

1. $T_{0,65}^{0} = \{1,7,8\}$
   $T_{0,65}^{1,2} = \{3,9,10\}$

   $\left|T_{0,65}^{0}\right| + \left|T_{0,65}^{1,2}\right| = 3 + 3 = 6 \geq 5$.

2. $T_{0,65}^{0,1} = \{1,2,4,5,6,7,8\}$
   $T_{0,65}^{2} = \{3,9\}$

$$\left|T_{0,65}^{0,1}\right| + \left|T_{0,65}^{2}\right| = 7 + 2 = 9 \geq 5 \cdot$$

Gauname, kad $\alpha_6 = 0{,}65$ taip pat – tinkama reikšmė. Visos kitos α reikšmės tikrinamos analogiškai.

Dabar sudarykime matricą A_2, kuri nuo matricos A_1 skiriasi tuo, kad veiksmai bus atliekami su trečiuoju matricos A stulpeliu – standartizuota gatvių tankio reikšme.

$$A_2^T = \begin{pmatrix} 8 & 1 & 6 & 7 & 4 & 5 & 2 & 9 & 10 & 3 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 2 & 1 & 2 \\ 0{,}2 & 0{,}3 & 0{,}3 & 0{,}3 & 0{,}4 & 0{,}6 & 0{,}7 & 0{,}7 & 0{,}7 & 0{,}8 \end{pmatrix} \cdot$$

Visos reikšmės trečiame stulpelyje išdėstomos didėjimo tvarka, o pirmas ir antras stulpeliai atitinkamai pasikeičia. Taikant (16) formulę nustatomos leistinos β reikšmės, t. y. $\beta_j \in \{0{,}25;0{,}35;0{,}50;0{,}65;0{,}75\}$. Parametro β tinkamumas parodytas su dviem reikšmėmis $\beta_2 = 0{,}35$ ir $\beta_2 = 0{,}65$.

1.
$$T_{0,35}^{0} = \{1,7,8\}$$
$$T_{0,35}^{1,2} = \{2,3,4,5,9,10\}$$
$$\left|T_{0,35}^{0}\right| + \left|T_{0,35}^{1,2}\right| = 3 + 6 = 9 \geq 5 \cdot$$

2.
$$T_{0,35}^{0,1} = \{1,6,7,8\}$$
$$T_{0,35}^{2} = \{3,9\}$$
$$\left|T_{0,35}^{0,1}\right| + \left|T_{0,35}^{2}\right| = 4 + 2 = 6 \geq 5 \cdot$$

Gauname, kad $\beta_2 = 0{,}35$ – tinkama reikšmė.

1.
$$T_{0,65}^{0} = \{1,7,8\}$$
$$T_{0,65}^{1,2} = \{2,3,9,10\}$$
$$\left|T_{0,65}^{0}\right| + \left|T_{0,65}^{1,2}\right| = 3 + 4 = 7 \geq 5 \cdot$$

2.
$$T_{0,65}^{0,1} = \{1,4,5,6,7,8\}$$
$$T_{0,65}^{2} = \{3,9\}$$

$$\left|T_{0,65}^{0,1}\right| + \left|T_{0,65}^{2}\right| = 6 + 2 = 8 \geq 5\,.$$

Gauname, kad $\beta_4 = 0{,}65$ – tinkama reikšmė. Visi kiti skaičiavimai atliekami analogiškai.

5. Algoritmas

Pasirinkdami atitinkamus α ir β iš matricos A gauname transformuotą matricą

$$T_{\alpha;\beta;\gamma\,...} : A \xrightarrow[\alpha,\beta;\gamma]{} T$$

$$T_{\alpha;\beta;\gamma,...} = \begin{pmatrix} k_1 & [x_1]_\alpha & [y_1]_\beta \\ k_2 & [x_2]_\alpha & [y_2]_\beta \\ ... & ... & ... \\ k_n & [x_n]_\alpha & [y_n]_\beta \end{pmatrix}; \qquad (17)$$

čia antrame ir trečiame stulpeliuose yra tik nuliai ir vienetai

$$[x_j]_\alpha = \begin{cases} 0, & x_j < \alpha, \\ 1, & x_j \geq \alpha, \end{cases}$$

$$[y_j]_\beta = \begin{cases} 0, & y_j < \beta, \\ 1, & y_j \geq \beta, \end{cases}$$

Grįžtant prie mūsų pavyzdžio, naujos matricos $T_{0,35,0,65}$ pirmame stulpelyje turime klasterizaciją pagal kainą, antrame $[x_j]_\alpha$ reikšmes, trečiame $[y_j]_\beta$ reikšmes.

$$T_{0,35;0,65}^{T} = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Bendruoju atveju turime tokią naujos klasterizacijos matricą:

$$K_{\alpha,\beta;\gamma,...} = \begin{pmatrix} k_1 & \tilde{k}_1 \\ k_2 & \tilde{k}_2 \\ ... & ... \\ k_n & \tilde{k}_n \end{pmatrix}. \qquad (18)$$

Mūsų pavyzdyje pirmame stulpelyje yra sena klasterizacija, o antrame – naujų klasterių numeriai, gauti sudėjus matricos $T_{0,35,0,65}$ antrojo ir trečiojo stulpelių elementus.

$$K^T_{0,35;0,65} = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 2 & 1 \\ 0 & 2 & 2 & 1 & 1 & 1 & 0 & 0 & 2 & 2 \end{pmatrix}.$$

Taigi pirmame stulpelyje gauname seną klasterizaciją pagal standartizuotą kainą, antrame stulpelyje naują klasterizaciją pagal du kintamuosius x ir y su atitinkamais dichotomizacijos slenksčiais $\alpha = 0{,}35$ ir $\beta = 0{,}65$. Paskutinis žingsnis – išmatuoti naujos klasterizacijos nesuderinamumo lygį pagal (5) formulę:

$$S(I(A), K) = 0 \cdot |T_0| + 1 \cdot |T_1| + 2|T_2| = 0 \cdot 8 + 1 \cdot 2 + 0 \cdot 2 = 2. \tag{19}$$

Geriausią klasterizaciją gausime parinkę α ir β reikšmes, minimizuojančias nesuderinamumo laipsnį $S(I(A), K)$.

Dabar pateiksime idealios klasterizacijos pavyzdį. Nagrinėkime matricą A

$$A^T = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ 0{,}1 & 0{,}1 & 0{,}1 & 0{,}1 & 0{,}1 & 0{,}7 & 0{,}7 & 0{,}7 & 0{,}7 & 0{,}7 \\ 0{,}2 & 0{,}2 & 0{,}2 & 0{,}8 & 0{,}8 & 0{,}2 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \end{pmatrix}.$$

Patikrinkime pasirinktų pagal (16) formulę reikšmių α ir β tinkamumą.

$\alpha = 0{,}4.$

$$\begin{matrix} \left|T^{0}_{0,4}\right| = 3, & \left|T^{1,2}_{0,4}\right| = 5 & \text{ir} & 3+5>5 \\ \left|T^{0,1}_{0,4}\right| = 5, & \left|T^{2}_{0,4}\right| = 4 & \text{ir} & 5+4>5 \end{matrix}.$$

$\beta = 0{,}5.$

$$\begin{matrix} \left|T^{0}_{0,5}\right| = 3, & \left|T^{1,2}_{0,5}\right| = 6 & \text{ir} & 3+6>5 \\ \left|T^{0,1}_{0,4}\right| = 4, & \left|T^{2}_{0,5}\right| = 4 & \text{ir} & 4+4>5 \end{matrix}.$$

$\alpha = 0{,}4$ ir $\beta = 0{,}5$ yra tinkamos reikšmės. Gauname matricą $T_{0,4;0,5}$, kurios pirmame stulpelyje yra pradinė klasterizacija. Sudėjus antro ir trečio stulpelio reikšmes, gaunama naujoji klasterizacija pagal du klausimus.

$$K^{T}_{0,4;0,5}=\begin{pmatrix}0&0&0&1&1&1&2&2&2&2\\0&0&0&0&0&1&1&1&1&1\\0&0&0&1&1&0&1&1&1&1\end{pmatrix}.$$

Matome, kad klasterizacjos visiškai sutampa. Taigi (5) funkcijos minimumo paieškos nuosekliai pagal kiekvieną p_i automatizavimas yra tolesnių tyrimų objektas.

Reikėtų ištirti, ar būtų tinkama godžioji strategija (Čiegis R., 2007).

6. Išvados

Šioje publikacijoje išnagrinėtas dichotominio diagnostinio testo matematinis modelis atskiruoju atveju, kuriame Vilniaus miesto mikrorajonai suskirstyti pagal įvertintą gyvenamojo nekilnojamojo turto kainą. Kuriamas naujas mikrorajonų suskirstymas pagal dichotomizuotus diagnostinius operatorius, kuris būtų artimiausias pradiniam suskirstymui pagal kainą. Artimiausiu laikomas suskirstymas, minimizuojantis klasterizacijų nesuderinamumo funkciją. Sukurtas algoritmas, leidžiantis geriausio dichotominio testo parinkimą atlikti automatizuotai.

Pateiktas klasterizacijos, sudarytos pagal trijų diagnostinių operatorių – mikrorajono atstumo nuo miesto centro, darbo vietų tankio ir gatvių tankio, sumos reikšmę pavyzdys. Taip pat pateiktas idealios klasterizacijos pagal du diagnostinius operatorius pavyzdys. Šiame straipsnyje detalizuotas metodas gali būti taikomas ir kitos rūšies nekilnojamojo turto apytikriam vertinimui.

Literatūra

An X, Deng Y, Fisher J.D. 2011. Commercial Real Estate Rental Index: A Dynamic Panel Data Model Estimation. A Research Paper Submitted to the Real Estate Research Institute (RERI).

Antuchevičienė, J. , Zakarevičius, A. , Zavadskas. 2011. E. *Measuring congruence of ranking results applying particular MCDM methods*. Informatica. 22(3): 319–338.

Chan, N. 2001 Stigma and its Assessment Methods.

Čiegis R. 2007. *Duomenų struktūros, algoritmai ir jų analizė*. Vilnius: Technika.

Krylovas A., Kosareva N. 2010. *Politominio diagnostinio testo matematinis modelis*. Liet. mat. rink. LMD darbai, 51: 279–284.

Krylovas A., Kosareva N. 2009. *Mathematical Modelling of Diagnostic Tests, Knowledge-Based Technologies and OR Methodologies for Strategic Decisions of Sustainable Development* (KORSD-2009): 5th international conference: 120-125.

Krylovas A., Kosareva N. 2008. *Mathematical modelling of forecasting the results of knowledge testing*. Technological and Economic Development of Economy 14(3): 388-401.

Krylovas A., Kosareva N., 2010. *Socialinių indikatorių konstravimas*. Social Technologies'10: challenges, opportunities, solutions. Conference Proceedings. Vilnius, November 25-26: 48-55.

Krylovas A., Kosareva N., Gudelytė L. 2011. *Socialinių indikatorių konstravimas taikant informacijos matavimo principus. Nekilnojamojo turto kainos modeliavimo pavyzdys*, Liet. mat. rink. LMD darbai, 52:195–199.

Krylovas A. 2009. *Diskrečioji matematika*. Vilnius: Technika.

Rudzkienė V., Burinskienė M. 2006. *Darnaus vystymosi principai Vilniaus miesto plėtroje*. Darnaus vystymosi strategija ir praktika: mokslo darbai. Vilnius: Mykolo Romerio universitetas, 39–46.

Scott R. Muldavin. 2010. Value beyond cost savings: how to underwrite sustainable properties, CRE, FRICS.

Lietuvos Respublikos turto ir verslo vertinimo pagrindų įstatymas. *Valstybės žinios*. 1999, Nr. 52-1672.

MODELING OF REAL ESTATE PRICES USING THE METHODICS OF DICHOTOMOUS TEST

Aleksandras Krylovas

Mykolas Romeris University, Lithuania, krylovas@mruni.eu

Natalja Kosareva

Vilnius Gediminas Technical University, Lithuania, natalja.kosareva@vgtu.lt

Laura Gudelytė

Mykolas Romeris University, Lithuania, l.gudelyte@mruni.eu

Tadas Laukevičius

Mykolas Romeris University, Lithuania, tadas@mruni.eu

Summary. *This article presents an analysis of the generalized dichotomous test methodologies. More as compared to the previous amount of available statistical information is used in this case. In addition, the authors seek to develop an algorithm which automatically calculates and clusters statistical data. As an example the authors take the neighbourhoods in the city of Vilnius . New algorithms will accurately cluster the Vilnius neighbourhoods by normalized price.*

Keywords: *mathematical modelling, creation of indicators, test, statistical methods.*