

КАДАСТР І МОНІТОРИНГ ЗЕМЕЛЬ

CADASTRE AND LAND MONITORING

УДК 332.3:528.44

Ю. ГУБАР

Кафедра кадастру територій, Національний університет "Львівська політехніка", вул. С. Бандери, 12, Львів, Україна, 79013, тел. +38(032)2582631, ел. пошта Yuriy.P.Hubar@lpnu.ua

ВПРОВАДЖЕННЯ МЕТОДИКИ КЛАСТЕРИЗАЦІЇ ДЛЯ ПОБУДОВИ ЕКОНОМІКО-МАТЕМАТИЧНИХ РЕГРЕСИВНИХ МОДЕЛЕЙ ОЦІНКИ НЕРУХОМОСТІ

Мета. Актуальною є проблема використання методики кластеризації для побудови економіко-математичних регресивних моделей оцінки нерухомості. Метою дослідження є побудова економіко-математичних моделей методами регресивного аналізу, що дасть змогу встановити не тільки найкращу оціночну модель для цієї конкретної вибірки, але й встановити принципи моделювання вартості об'єктів нерухомості на основі відомих цін аналогів. Реалізацію алгоритму регресивного аналізу необхідно розпочати з побудови матриці вибірових парних коефіцієнтів кореляції. Обчислення та аналіз коефіцієнтів кореляції дозволять встановити зв'язок вартості з кожним із чинників [МСО-1, 2006; МСО-2, 2006; НС-1, 2003; НС-2, 2004]. **Методика.** Методика ґрунтується на застосуванні методу дендрограм (алгоритми алгомеративно-ієрархічного кластерного аналізу) з використанням просторово стиснених чинників (лінійних комбінацій вихідних чинників). З метою вибору методу групування об'єктів нерухомості доцільно провести порівняльний аналіз та тестування різних методів кластерного аналізу. **Результати.** Виконані дослідження доводять важливість застосування методики кластеризації для впровадження в Україні масової оцінки нерухомості. Якість економіко-математичних моделей визначають характеристики адекватності, стійкості та корисності моделі, які можна трактувати як погодження інформації, що описує функціональні можливості моделі. **Наукова новизна.** Виконані дослідження дають змогу отримувати кінцеві оціночні моделі для оцінювання неочінованих об'єктів будь-якого кластеру. Побудова моделі передбачає ретельне збирання ринкової інформації і детальне аналітичне опрацювання. Під час побудови моделі з метою отримання оптимального варіанта моделі, що дасть змогу уникнути недовраховання наявних чинників та не допустить включення в модель надто великої кількості змінних, пропонується зменшити кількість чинників з використанням багатокрокового регресивного аналізу. **Практична значущість.** У роботі виконано практичну реалізацію теоретичних розробок. Здійснено побудову оціночних моделей методами регресивного аналізу і перевірку якості оцінки на прикладі відібраного кластеру (садибна забудова) Львівського регіону. Представлене в роботі дослідження доцільно застосовувати для населених пунктів України [Губар Ю., 2012].

Ключові слова: оцінка нерухомості, економіко-математичне моделювання, методика кластеризації, регресивні моделі.

Вступ

Для побудови як комплексу взаємозв'язаних економіко-математичних моделей (ЕММ), так і будь-якої окремої моделі, необхідна певна сукупність принципів, які дають можливість коректно формалізувати моделюючі системи та об'єкти. Як загальні принципи економіко-математичного моделювання доцільно застосувати принципи: системності, інтегрованості, невизначеності головних видів діяльності, достатності, інваріантності, наступності та ефективності.

Інтенсивний розвиток досліджень проблематики побудови економіко-математичних моделей для оцінювання земель та нерухомості розпочався понад століття тому. Розроблено оціночні моделі у вигляді простих алгебраїчних формул, однак велика кількість спрощень закладених у цих моделях не давала змогу отримувати достовірні результати. Національну асоціацію податкових оцінювачів, а в подальшому Міжнародну асоціацію податкових оцінювачів (IAAO) створено у США в 1934 році з метою удосконалення методів, моделей і процедур

оцінки нерухомості шляхом підвищення кваліфікації оцінювачів, здійснення досліджень та забезпечення уніфікації практичних процедур оцінки. Розроблено великий спектр оціночних моделей, заснованих на економіко-математичному моделюванні. Економіко-математичні моделі оцінки нерухомості постійно вдосконалюються. З розвитком комп'ютерної техніки стало можливим опрацювання великих масивів даних, що має особливе значення для кадастрової оцінки нерухомості, оскільки здійснюється оцінка великої кількості об'єктів та виникає необхідність аналізу великого об'єму інформації [Фрідман Дж., 2012; Фрідман Дж., 2013].

Мета

Актуальною є проблема використання методу кластеризації для побудови економіко-математичних регресивних моделей оцінки нерухомості. Метою дослідження є побудова економіко-математичних моделей методами регресивного аналізу, що дасть змогу встановити не тільки найкращу оціночну модель для конкретної вибірки, але й встановити принципи моделювання вартості об'єктів нерухомості на основі відомих цін аналогів. Реалізацію алгоритму регресивного аналізу необхідно розпочати з побудови матриці вибірових парних коефіцієнтів кореляції. Обчислення та аналіз коефіцієнтів кореляції дасть змогу встановити зв'язок вартості з кожним із чинників [МСО-1, 2006; МСО-2, 2006; НС-1, 2003; НС-2, 2004].

Методика

Методика ґрунтується на застосуванні методу дендрограм (алгоритми алгомеративно-ієрархічного кластерного аналізу) з використанням просторово стиснених чинників (лінійних комбінацій вихідних чинників). З метою вибору методу групування об'єктів нерухомості доцільно провести порівняльний аналіз та тестування різних методів кластерного аналізу. Під час побудови оцінювальних моделей (адитивних, мультиплікативних) усередині кожного кластера необхідно здійснювати аналіз з погляду найадекватнішого відображення реально-го ринку нерухомості і для цього необхідно

сформувані принципи вибору та умови застосування різних варіантів оціночних моделей.

Ефективність прийняття рішень великою мірою залежить від рівня досягнутої мети дослідження, яка, своєю чергою, визначається цільовою функцією. Моделі можуть будувати для досягнення певних цілей. Розглянемо їх:

1. Виявлення функціональних співвідношень – визначення кількісних залежностей між вхідними факторами моделі та вихідними характеристиками об'єкта дослідження. Подібні моделі за своїм характером є описовими і присутні під час побудови математичних моделей будь-яких видів.

2. Аналіз чутливості – встановлення з великої кількості наявних факторів тих, які значною мірою впливають на вихідні характеристики об'єкта дослідження. Моделі аналізу чутливості повинні обов'язково передбачати можливість варіації низки факторів і їх так само можна використати для оцінювання точності розв'язків, отриманих за допомогою моделей будь-яких типів.

3. Прогноз – оцінка поведінки об'єкта на часовому інтервалі за деякого допустимого послідовного зовнішніх умов. Переважно задачі прогнозу є динамічними відносно вхідних параметрів і як незалежна змінна виступає час. Моделі прогнозу також є описовими.

4. Оцінка – визначення, наскільки адекватно об'єкт дослідження відповідає деяким критеріям. На відміну від розглянутих вище типів моделей, моделі оцінки включають розрахунки інтегральних характеристик-критеріїв, які формалізують мету дослідження. Моделі оцінки займають проміжне місце між описовими й оптимізаційними моделями. У них задані критерії і їх деяке "критичне" значення, але проводиться не оптимізація, а лише порівняння розрахункового значення з "критичним", після чого приймається рішення про задоволення характеристик об'єкта поставленим вимогам.

5. Порівняння – зіставлення обмеженого числа альтернативних варіантів систем або зіставлення декількох прийнятних принципів чи методів дій. Кількість варіантів припускається незначним, у зв'язку з чим оцінюються всі варіанти, тобто здійснюється прямий перебір

всієї множини. Хоча моделей окремого типу, близьких до оптимізаційних, спеціального блоку оптимізації у них немає.

6. Оптимізація – точне визначення такого послання змінних керування, за якого забезпечується екстремальне (максимальне або мінімальне залежно від змісту критерію оптимальності) значення цільової функції. Суттєва різниця від наведеного вище випадку – наявність спеціального блоку оптимізації, який дає змогу цілеспрямовано і найефективніше з обчислювального погляду переглядати множину альтернативних варіантів [Харрисон Г., 1994; Еккерт Д., 1997; Еккерт Д., 1999].

Існують сім видів економіко-математичних моделей, які можна застосовувати для оцінювання нерухомості, а саме:

1. Регресивні адитивні лінійні.

$$Y = a_0 + a_1 \cdot x_1 + \dots + a_n \cdot x_n, \quad (1)$$

де Y – питома вартість об'єкта; x_1, \dots, x_n – чинники вартості; a_0, a_1, \dots, a_n – коефіцієнти моделі.

2. Регресивні мультиплікативні: степенева, експонентна.

$$Y = a_0 \cdot x_1^{a_1} \cdot \dots \cdot x_n^{a_n}, \quad (2)$$

де a_0, a_1, \dots, a_n – ваги чинників.

3. Регресивні адитивні нелінійні, але лінійні за коефіцієнтами.

$$Y = a_0 + a_1 \cdot \varphi_1 \cdot (S_1) + \dots + a_n \cdot \varphi_n \cdot (S_n), \quad (3)$$

де S_1, \dots, S_n – задані сукупності чинників; $\varphi_1 \cdot (S_1) \dots \varphi_n \cdot (S_n)$ – задані нелінійні функції від чинників.

4. Регресивні гібридні.

$$Y = x_1^{B_1} \cdot \dots \cdot x_n^{B_n} \cdot (A_0 + A_1 \cdot x_1 + \dots + A_n \cdot x_n), \quad (4)$$

де A, B – ваги чинників.

5. Просторово-параметричні – добуток множників коригування на середню вартість нерухомості із вибірки.

$$Y_{ind} = \bar{x} \times k_{iq} \times k_{nB} \times k_{dA}, \quad (5)$$

де $k_{iq} = \frac{\bar{x}_{iq}}{\bar{x}}$ – коефіцієнт місцерозташування і т.д.

6. Емпіричні.

$$Y = \frac{Y_{cp}(NM)}{Q(NM)} \cdot \Phi(h_i, x_i, i=1, \dots, m), \quad (6)$$

де $Y_{cp}(NM)$ – середня ціна об'єктів із NM;

x_i – суттєві чинники; $h_i = \frac{\bar{x}_{NM}}{\bar{x}}$, Φ – функція

заданого виду від добутоків h_i та x_i ; $Q(NM)$ – функція того самого виду від середніх значень.

7. Нейронні мережі представлені декількома шарами проміжних нейронів.

У табл. 1 подано аналіз наявних моделей оцінки нерухомості [Харрисон Г., 1994; Еккерт Д., 1997; Еккерт Д., 1999].

На нашу думку, для виконання розрахунків у кадастровій оцінці нерухомості доцільно обмежитися використанням регресивних моделей, тому що якість оцінки цих моделей не поступається якості оцінки моделей інших класів. Ці висновки підтверджені вивченням результатів апробації різноманітних економіко-математичних моделей оцінки нерухомості експертів-оцінювачів України.

У табл. 2 подано аналіз методик кадастрової вартості нерухомості [Харрисон Г., 1994; Еккерт Д., 1997; Еккерт Д., 1999].

Гібридна методика потребує застосування спеціального програмного забезпечення та високої кваліфікації виконавців з математичної статистики та методів мінімізації складних функцій. Це складно забезпечити з урахуванням стану обчислювальної техніки та освіти персоналу, але основним обмеженням застосування гібридних моделей є нерозвиненість ринку нерухомості. Суттєвими недоліками локального підходу є неможливість аналізу локальних моделей традиційними статистичними методами.

Найперспективнішою для виконання кадастрової оцінки нерухомості населених пунктів України, на нашу думку, є методика з використанням кластеризації, яка передбачає таку послідовність дій:

1) збір статистичних даних про об'єкти оцінки;

2) визначення складу ціноутворювальних чинників;

3) попереднє опрацювання вихідної вибірки, а саме: видалення неповних описів та нетипових даних;

4) формування первинних змінних на основі чинників;

Таблиця 1

Аналіз моделей оцінки нерухомості

Table 1

Analysis of real estate models evaluation

Вид моделей	Сфера застосування
Регресивні адитивні лінійні	<ol style="list-style-type: none"> 1. Сукупність значень чинників і вартостей допускає представлення у вигляді багатомірного нормального розподілу. 2. Похибки моделі однаково розподіляються, мають нульове середнє значення і однакові середні квадратичні помилки. 3. Вплив кожного чинника на вартість не залежить від його значень, тобто коефіцієнти моделі вважаються постійними. 4. Добрі результати оцінки можна отримати лише для великих сукупностей об'єктів з близькими значеннями чинників. 5. Обов'язкова процедура кластеризації. 6. Оптимальна для ринку, який достатньо забезпечений інформацією. 7. Критерій якості: коефіцієнт детермінації R^2, середня відносна похибка.
Регресивні мультиплікативні	<ol style="list-style-type: none"> 1. Перехід до лінійної моделі шляхом логарифмування. 2. Дають змогу коректніше відобразити залежність вартості від чинників. 3. Не потребують дотримання умов адитивності. 4. Добрі результати оцінки можна отримати лише для великих сукупностей об'єктів з близькими значеннями чинників. 5. Обов'язкова процедура кластеризації. 6. Оптимальна для ринку, який достатньо забезпечений інформацією. 7. Критерій якості: коефіцієнт детермінації R^2, середня відносна похибка.
Регресивні адитивні нелінійні, але лінійні за коефіцієнтами	<ol style="list-style-type: none"> 1. Пункти 1–3 представлені для коректного застосування лінійних моделей повинні виконуватись для значень нелінійних функцій від чинників вартості. 2. Необхідний аналіз власних значень коваріаційної матриці. 3. Коректні результати оцінки можна отримати лише за великої кількості власних значень коваріаційної матриці. 4. Критерій якості: коефіцієнт детермінації R^2, середня відносна похибка.
Регресивні гібридні	<ol style="list-style-type: none"> 1. Відсутня обмежувальна умова для вихідної вибірки. 2. Модель має відповідну канонічну форму. 3. Оптимальна для ринку із встановленими тенденціями та складними видами залежностей. 4. Залежність результату від початкових значень.
Просторово-параметричні	<ol style="list-style-type: none"> 1. Основна частка вартості об'єктів визначається невеликою множиною значущих чинників. 2. Необхідна перевірка статистичної значущості середніх значень та вектору цих значень. 3. Модель проходила апробацію на розвинутих ринках, але якість оцінки значно знижується під час переходу до недостатньо розвинутих ринків. 4. Під час застосування моделі для неоднорідних вибірок виникають методичні проблеми.
Емпіричні	<ol style="list-style-type: none"> 1. Вигляд моделі залежить від структури кореляційних зв'язків між чинниками і розкладом нелінійної функції за формулою Тейлора. 2. Необхідна перевірка значущості середніх значень та вектора цих значень. 3. Не вимагає застосування алгоритмів мінімізації. 4. Цей клас моделей знаходиться в стадії розроблення.
Нейронні мережі	<ol style="list-style-type: none"> 1. Вибірка повинна бути досить великою та унімодальною. 2. Якість оцінки залежить від вибору початкових умов і вміння розпізнавати локальні мінімуми. 3. Якість перевіряється за значеннями цільової функції. 4. Для кожного району потрібно мати окрему нейронну мережу, або адаптувати базову.

Таблиця 2

Аналіз методик кадастрової вартості нерухомості

Table 2

Вид методики	Методика з використанням кластеризації	Методика з використанням локальних моделей	Гібридна методика
Основні етапи	<p>Методика з використанням кластеризації</p> <ul style="list-style-type: none"> • розподіл вихідних даних на дві вибірки, одна з яких містить тільки забудовані земельні ділянки, а інша – незабудовані (кожна вибірка розглядається окремо); • первинне групування кожної із вибірок за адміністративно-територіальними ознаками та їх функціонального використання; • для кожної із вибірок здійснюється вибір сукупностей чинників; • видалення неповних описів в кожній вибірці, перевірка достовірності даних; • видалення описів з рідкими стрічками (аномальними) ринкових цін; • формування первинних змінних моделі на основі ціноутворювальних чинників; • формування таблиці опису ділянок у розрізі первинних змінних; • кластеризація: <ul style="list-style-type: none"> а) первинна – впорядкування рядків вихідної таблиці за значеннями одного чинника; б) визначення сукупності чинників вторинної кластеризації; в) вторинна кластеризація за принципом близькості їх опису за чинниками вартості, яка здійснюється з використанням статистичного методу головних компонент. 	<p>Методика з використанням локальних моделей</p> <ul style="list-style-type: none"> • розподіл вихідних даних на дві вибірки, одна з яких містить тільки забудовані земельні ділянки, а інша – незабудовані (кожна вибірка розглядається окремо); • первинне групування кожної із вибірок за адміністративно-територіальними ознаками та їх функціонального використання; • для кожної із вибірок здійснюється вибір сукупностей чинників; • видалення неповних описів в кожній вибірці, перевірка достовірності даних; • видалення описів з рідкими стрічками (аномальними) ринкових цін; • формування первинних змінних моделі на основі ціноутворювальних чинників; • визначення значущих чинників (первинні значущі чинники); • значущі чинники з неперервною множиною значень розбиваються на інтервали, в середині яких ці значення вважаються близькими; • формування коду об'єктів оцінки, елементами якого є або відповідні номери інтервалів розбиття або відповідні значення значущих чинників; • вихідна таблиця розбивається на сукупність клітин з різними кодами (кліткова структура вихідної вибірки); • визначення сукупності значущих чинників всередині кожної клітини (визначення вторинних значущих чинників). 	<p>Гібридна методика</p> <ul style="list-style-type: none"> • формування бази даних за ринковою інформацією; • рекурентна процедура вибору чинників вартості, що суттєво впливають на вартість об'єктів нерухомості з використанням гібридної моделі спеціального виду; • процедура покрокової регресії, де виконується коригування переліку чинників згідно з ступенем їх значущості та відкидання чинників з малою вагою; • з використанням картографічної інформації розраховують чинник географічного впливу та вводиться в модель.

5) формування таблиці опису об'єктів у розрізі змінних;

6) кластеризація (групування) об'єктів оцінювання;

7) побудова регресивних оціночних моделей;

8) контроль якості оцінювання;

9) розрахунок вартості неоцінених об'єктів.

Виконаний нами детальний аналіз основних етапів цієї методики дав змогу виявити основні проблеми процесу оцінки, а саме:

- на другому етапі відсутній механізм відбору ціноутворювальних чинників та присутній вплив суб'єктивізму оцінювача;

- на третьому етапі методично невірно автоматично приймається гіпотеза нормальності розподілу значень;

- на четвертому етапі методично неправильно здійснюється кодування якісних характеристик нерухомості;

- на шостому етапі під час групування об'єктів методом дендрограм цінова інформація використовується не повністю і внаслідок використання лінійних комбінацій вихідних показників ускладнена інтерпретація результатів кластеризації;

- на сьомому етапі – вид моделі вибирається наперед, що на практиці надскладно і внаслідок використання лінійних комбінацій вихідних показників складно інтерпретувати результати моделювання;

- на восьмому етапі немає поділу вибірки на навчальну і перевіряючу, немає перевірки нормальності розподілів залишків, немає перевірки статистичної значущості рівняння за F-чинником, немає перевірки значущості коефіцієнтів за t-статистикою, немає аналізу мультиколінеарності чинників, структури похибок.

Для відбору ціноутворювальних чинників необхідно залучати досвідчених спеціалістів відповідного ринку нерухомості та розробити механізм формального вибору чинників.

На етапі попереднього опрацювання вихідної вибірки необхідно автоматично приймати гіпотезу про нормальність розподілу вибірки та виконати відсів нетипових значень вибірки. Він полягає у видаленні значень, що відрізняються від середнього більше ніж на 2 стандартних

відхилення [Перович Л., 2010; Губар Ю., 2013; Губар Ю., 2013; Губар Ю., 2012].

Основною проблемою етапу “формування первинних змінних моделі на основі ціноутворювальних чинників” є врахування в оцінювальних моделях якісних чинників. Необхідно провести порівняльний аналіз та апробацію різних методів оцифрування якісних ознак, що дасть змогу встановити найвпливовіші.

Для кластеризації (групування) об'єктів оцінки необхідно застосовувати метод дендрограм (алгоритми алгомеративно-ієрархічного кластерного аналізу) з використанням просторово стиснених чинників (лінійних комбінацій вихідних чинників). З метою вибору методу групування об'єктів нерухомості доцільно провести порівняльний аналіз та тестування різних методів кластерного аналізу.

Під час побудови оціночних моделей (адитивних, мультиплікативних) всередині кожного кластера необхідно здійснювати з погляду найадекватнішого відображення реального ринку нерухомості і для цього необхідно сформулювати принципи вибору та умови застосування різних варіантів оціночних моделей.

Для того, щоб отримати оптимальний варіант моделі, який дасть змогу уникнути недобрахування наявних чинників та не допустить включення в модель надто великої кількості змінних, необхідно зменшити кількість чинників з використанням багатокрокового регресивного аналізу під час побудови моделі.

Для визначення якості оцінювальних моделей необхідно використати:

- графік різниць між відомими питомими ринковими цінами і їх модельними аналогами та суму квадратів різниць;

- коефіцієнт детермінації R^2 ;

- середню відносну похибку;

- для перевірки якості отриманої моделі використати дані, що не використовувались для побудови моделі;

- перевірку значущості побудованого рівняння;

- перевірку значущості коефіцієнтів при ціноутворювальних чинниках;

- оцінку ступеня впливу кожного чинника на вартість об'єкта оцінки;

- аналіз мультиколінеарності ціноутворювальних чинників.

Отже, в результаті досліджень системи кадастрової оцінки нерухомості, вивчення закордонного досвіду у сфері оцінки земель та нерухомості, аналізу праць вітчизняних та закордонних спеціалістів найперспективнішим, на наш погляд, є методики, засновані на економіко-математичному моделюванні.

Виконано порівняльний аналіз економіко-математичних моделей, який показав, що для розрахунку кадастрової вартості нерухомості доцільно обмежитися використанням регресивних моделей, тому що незважаючи на їх відносну простоту, за якістю оцінки вони не поступаються моделям інших класів. Зі всіх методів регресивного аналізу найперспективнішою, на наш погляд, є методика з використанням кластеризації.

Особливості відбору ціноутворювальних чинників та їх врахування в оцінювальних моделях.

Оцифровування кількісних і якісних ознак

Для відбору ціноутворювальних чинників необхідно оцінити їх ступінь впливу на питому вартість через обчислення часткового коефіцієнта кореляції між цими чинниками і вектором питомих вартостей, а також проаналізувати мультиколінійність чинників шляхом обчислення парного коефіцієнта кореляції кожної пари чинників. Існують дві категорії залежностей (функціональна і кореляційна) і дві групи ознак (ознаки-чинники і результуючі ознаки). На відміну від функціонального зв'язку, де існує стовідсоткова відповідність між чинниками (X) і результативними ознаками (Y), у кореляційному зв'язку немає повної відповідності. Змінні можуть вимірюватися різними шкалами і це визначає вибір відповідного коефіцієнта кореляції. Ми пропонуємо використовувати параметричний метод розрахунку коефіцієнта кореляції Пірсона r для кількісних чинників і непараметричний коефіцієнт кореляції рангів Спірмена r_s – для якісних чинників.

Відбір значущих чинників із типового їх переліку на основі кореляційного аналізу необхідно виконувати з такими допусками:

- вибірка повинна мати достатній об'єм (залежить від кількості чинників);

- вибірка за кожним чинником повинна бути однорідною, що забезпечує незміщеність оцінки середніх величин;

- матриця спостережень не повинна містити пропусків;

- розподіл значень повинен підпорядковуватись нормальному закону, оскільки необхідна перевірка значущості коефіцієнтів кореляції.

Алгоритм цього відбору повинен бути таким:

- 1) обчислення оцінок парних коефіцієнтів кореляції між ціноутворювальними чинниками. Перевірка значущості оцінок;

- 2) обчислення оцінок часткових коефіцієнтів кореляції між кожним чинником та вектором питомих цін. Перевірка значущості оцінок;

- 3) відбір усіх значущих чинників і створення списку ціноутворювальних чинників.

Для відбору ціноутворювальних чинників пріоритетними вважаються чинники з найменшим значенням парних коефіцієнтів кореляції.

Для врахування всіх чинників у регресивній моделі якісні чинники необхідно перевести до кількісного виду. Кількісна ознака – це ознака, значення якого отримане в результаті вимірювання, спостереження або розрахунку, виражається відповідним числом. Якісна ознака – це ознака, що характеризує деяку властивість або стан, а також наявність або відсутність цієї властивості в об'єкта.

Якісні характеристики поділяються на порядкові (існує можливість встановлення розбіжності або еквівалентності об'єктів за відповідною ознакою і розбиття їх на декілька класів) та номінальні (існує можливість визначення якісних переваг між об'єктами за інтенсивністю прояву цієї властивості). Для номінальних ознак всім об'єктам одного класу призначають однакові числові значення, для порядкових ознак значення задаються у балах і у рангах. Частковим випадком якісних ознак є бінарні змінні, які описують наявність або відсутність у об'єкта будь-якої якості.

Кожному типу ознак відповідає свій тип шкал вимірювань (кількісні, не кількісні – порядкові, номінальні), група допустимих перетворень значень шкали і підмножини коректних методів опрацювання величин. Поділ кількісних

ознак на неперервні та дискретні умовний, оскільки:

- через обмеження точності вимірів неперервні за своєю природою показники можна розглядати як дискретні;

- методи, що застосовуються для опрацювання неперервних кількісних даних, можна ефективно застосовувати для опрацювання дискретних ознак з невеликою кількістю градацій.

Теорія лінійних регресивних моделей із ненульовим вільним членом не встановлює жодних обмежень на характер значень кількісних ознак – вони можуть бути безперервними і дискретними, а також бінарними [The Appraisal of Real Estate. Twelfth Edition – Chicago: Appraisal Institute, 2011]. Для кількісних показників використовують математичні моделі неперервних величин так званий класичний апарат багатомірного статистичного аналізу. Для номінальних і порядкових шкал доцільні інші математичні моделі та методи [Перович Л., 2010; Губар Ю., 2008]. На практиці застосовують випадки поєднання різнотипних ознак. Для застосування кількісних методів регресивного аналізу допускаються такі перетворення вихідної інформації:

- номінальні ознаки зводяться до системи бінарних змінних;

- деякі порядкові ознаки можуть застосовуватися як квазікількісні, інші – зводяться до сукупності бінарних;

- всі ознаки зводяться до кількісного типу за допомогою оптимізаційних процедур.

У практичних завданнях оцінки нерухомості доцільне спільне використання кількісних, квазікількісних і сукупностей бінарних показників. Це дасть змогу найбільшою мірою врахувати статистичний зв'язок всіх ознак, звести до мінімуму втрату інформації під час перетворень і використати потужний апарат класичних методів аналізу числових даних.

Ознаки нечислової природи для врахування їх в регресивних моделях зводяться до числового виду шляхом оцифрування, тобто шляхом привласнення їх значенням деяких числових позначок. Розбивка на класи (для номі-

нальних та вимірних в балах порядкових ознак) рекомендується зі 3–7 градацією. Статистична процедура, що дає змогу обчислити необхідну кількість градацій залежно від діапазону допустимих кількісних змін ознаки і дисперсії помилок експертів під час визначення значення ознаки, наведено в [Перович Л., 2010; Губар Ю., 2008].

На рис. 1 подано підходи до оцифрування не кількісних ознак.

Наприклад, номінальна ознака “якість електрофікації” з градаціями “електрокабель”, “повітряна лінія”, “відсутність електрофікації” можна перетворити на порядкову шкалу через дослідження ринку нерухомості. Так, на ринку найбільше цінуються земельні ділянки, що забезпечені електрокабелем, найменше – на яких немає електрофікації, тому для градації номінальної ознаки можна ввести відношення порядку та перейти до порядкової ознаки в балах: 1 – відсутність електрофікації, 2 – повітряна лінія, 3 – електрокабель. Причому масштаб і початок шкали не суттєві. Не правильний поділ на класи і вибір кількості градацій можуть призвести до отримання абсурдних результатів.

Від вибраної процедури оцифрування залежить якість регресивної моделі. Бінарні ознаки можна оцифрувати довільно, але частіше їм присвоюють значення 0 і 1. У класичній лінійній регресивній моделі використовують таку залежність:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_jx_j + \dots + a_kx_k. \quad (7)$$

Для врахування не кількісної ознаки x_j з m градаціями $\{x_j^1, x_j^2, \dots, x_j^m\}$ потребує введення $m-1$ бінарної змінної: z_1, z_2, \dots, z_{m-1} .

Рівняння (7) можна записати у вигляді:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_{j-1}x_{j-1} + b_1z_1 + b_2z_2 + \dots + b_{m-1}z_{m-1} + a_{j+1}x_{j+1} + \dots + a_kx_k. \quad (8)$$

Для порядкових ознак, крім оцифрування на основі бінарних змінних, існують й інші способи їх врахування в оцінювальних моделях [The Appraisal of Real Estate. Twelfth Edition – Chicago: Appraisal Institute, 2011].

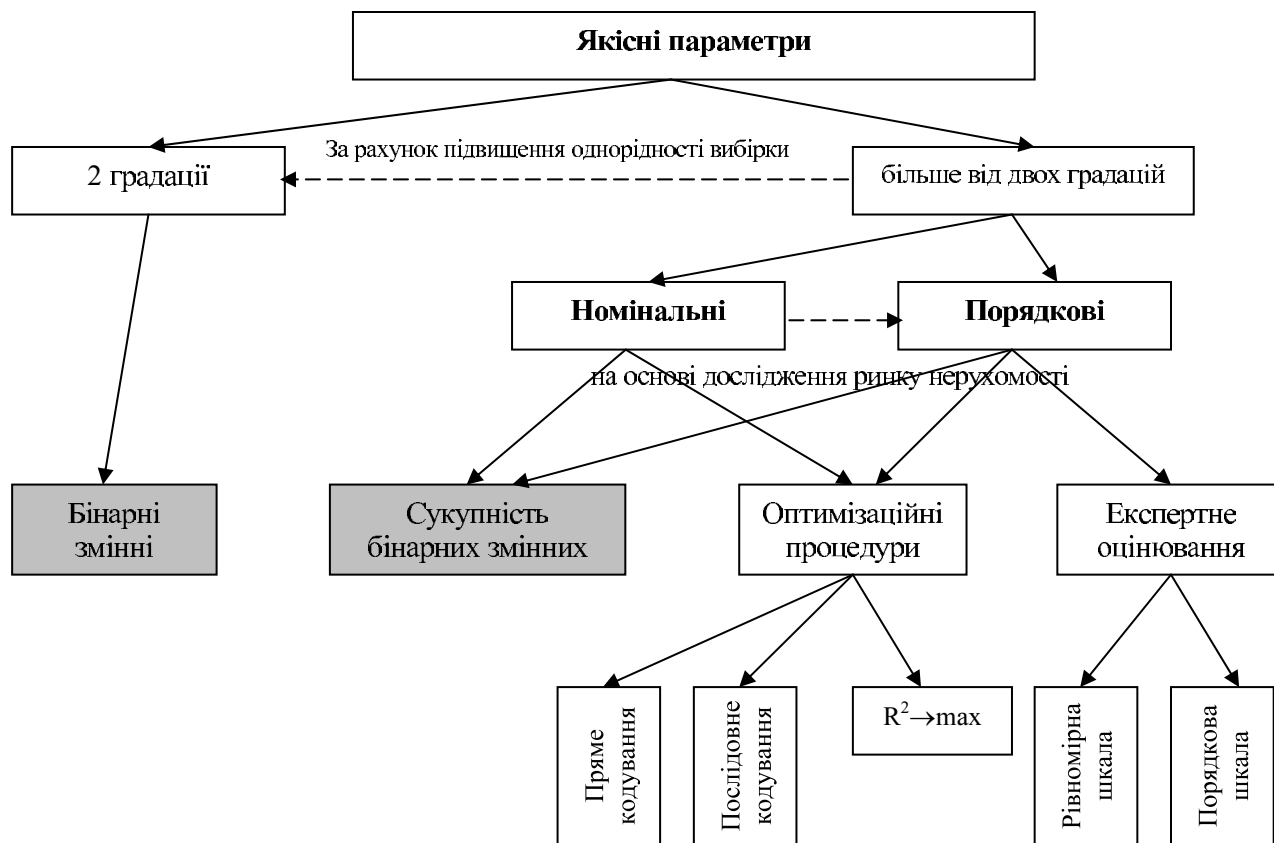


Рис. 1. Методи оцифрування якісних ціноутворювальних чинників
 Fig. 1. Methods of digitizing of qualitative pricing factors

Практична реалізація досліджень. Побудова оцінювальних моделей методами регресивного аналізу і перевірка якості оцінки (на прикладі Львівського регіону)

Виконаємо побудову економіко-математичних моделей методами множинного регресивного аналізу. Для дослідження виберемо найперспективніший з погляду подальшого розвитку сектор нерухомості, як землі садибної забудови Львівського регіону. Після дослідження та аналізу цього сегменту ринку нерухомості ми встановили 22 об'єкти, що найбільший кластер із усіх об'єктів цього сегменту. Дослідження виконаємо для 21 об'єкта, а показники 22 об'єкта будуть використані для перевірки отриманої моделі.

Ми встановили десять найзначущіших ціноутворювальних чинників, що впливають на вартість об'єкта оцінки, а саме:

- відстань до межі Львова, км – кількісний чинник;
- відстань до водойми, км – кількісний чинник;

- відстань до лісу, км – кількісний чинник;
- водопостачання – якісний чинник;
- газифікація – якісний чинник;
- електрифікація – якісний чинник;
- каналізація – якісний чинник;
- відстань до межі районного центру, км – кількісний чинник;
- відстань до з/д станції, км – кількісний чинник;
- розмір ділянки, сотки – кількісний чинник.

Для перевірки вихідних даних на нормальність варіаційного ряду з метою для побудови оцінювальних моделей сформуємо варіаційний ряд цін у розрахунку за 1 сотку земельної ділянки (табл. 3).

Отже, вибіркове середнє варіаційного ряду: $\bar{x} = 1468,83$, стандартне відхилення $s_n = 464,6629$.

Довірчі інтервали:

- 90 % – $1468,83 \pm 166,7845$;

Таблиця 3

Table 3

Вихідні дані для побудови оціночних моделей (виділений кластер)

Initial data for creation of value models (the selected cluster)

№ з/п	Відстань до межі Львова, км	Відстань до водойми, км	Відстань до лісу, км	Водопостачання	Газифікація	Електрифікація	Каналізація	Відстань до межі районного центру, км	Відстань до з/д станції, км	Розмір ділянки, сотки	Вартість ділянки, у.о.	Ціна за сотку, у.о.
1	15	0,2	0,2	0	1	1	0	2	0,3	14	30002	2143
2	12	0,9	2	0	1	0,5	0	3	3	25	52000	2080
3	12	0,8	1,2	0	1	0,5	1	1	3	20	44000	2200
4	14	0,2	0,7	0	1	0,5	1	3	2	12	26004	2167
5	15	0,2	1,8	0	1	0,5	1	4	2	15	31995	2133
6	20	1	1	0	1	1	0	2,5	1	27	45009	1667
7	22	0,5	1	0	0	1	0	4,5	0,7	10	14000	1400
8	25	0,3	0,5	0	1	1	1	6	0,2	15	22005	1467
9	35	0,4	1,4	0	1	1	0	10	1,5	11	9999	909
10	33	0,1	1,5	0	0	1	0	7	2	12	11004	917
11	25	0,3	0,2	0	1	1	1	6	0,2	15	24000	1600
12	22	0,5	1	0	1	1	0	4,5	0,7	8	13000	1625
13	20	1	1	0	0	1	0	2,5	3	6	7500	1250
14	17	0,6	1,2	0	1	1	0	1	0,2	24	43008	1792
15	35	1,5	0,2	0	1	1	0	2	3	9	8001	889
16	22	0,3	0,8	0	0	0,5	0	12,5	0,1	12	12000	1000
17	22	0,2	0,3	0	0	0,5	0	12,5	0,1	12	12996	1083
18	22	0,2	0,3	0	0	0,5	0	12,5	0,8	15	16500	1100
19	22	0,3	0,8	0	0	0,5	0	12,5	0,8	15	18000	1200
20	28	0,1	0,1	0	1	1	0	7	0,2	15	16005	1067
21	26	1,5	0,3	0	1	1	0	5	1	19	22002	1158
22	20	0,4	0,5	0	1	1	0	12,5	0,2	12	18000	1500

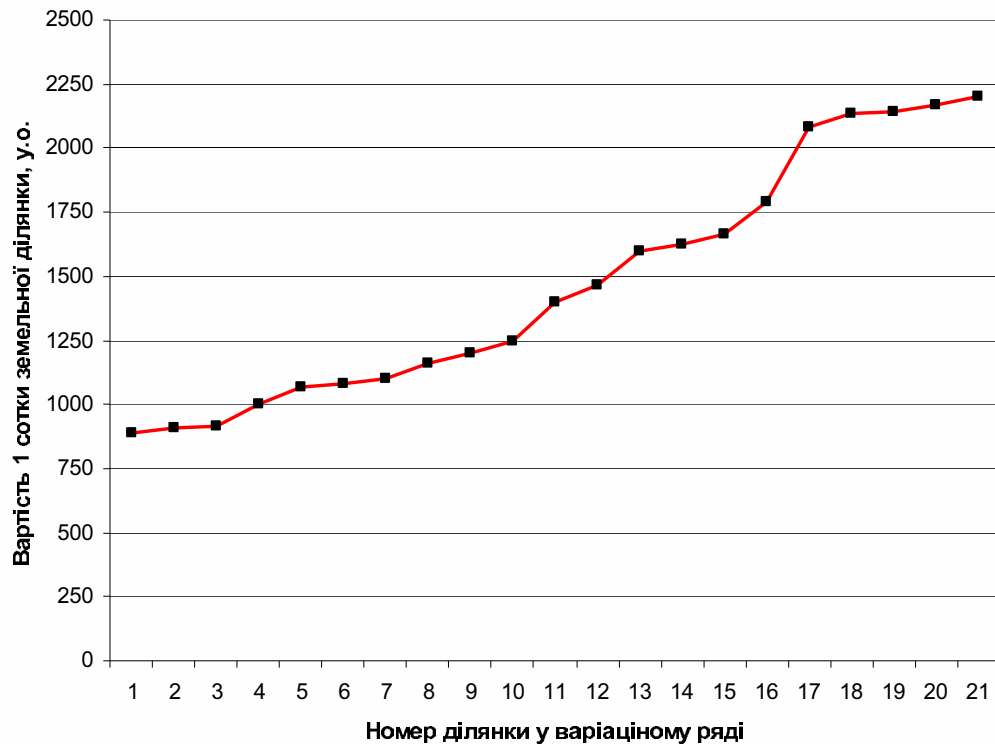


Рис. 2. Графік титомих вартостей виділеного кластеру

Fig. 2. Graph of the specific values of selected cluster

Таблиця 4

Перевірка вибірки на грубі похибки

Table 4

Перевірка вибірки на грубі похибки

Об'єм вибірки $n = 21$		Мінімальне значення 889	Максимальне значення 2200	
Критерій Смірнова–Грбса	Статистики		$T_{(1)} = 1,248103$	$T_{(n)} = 1,573550$
	Критичні значення	$C_{10\%} = 2,44$	Приймається	Приймається
		$C_{5\%} = 2,63$	Приймається	Приймається
		$C_{2,5\%} = 2,95$	Приймається	Приймається
Критерій Грбса	Статистики		$G_{(1)} = 0,91182$	$G_{(n)} = 0,870001$
	Критичні значення	$C_{10\%} = 0,44$	Приймається	Приймається
		$C_{5\%} = 0,37$	Приймається	Приймається
		$C_{2,5\%} = 0,31$	Приймається	Приймається
Критерій Тіг'сна–Мура			пара мін. значень 889, 909	пара макс. значень 2200, 2167
	Статистики		$L_{(2)} = 0,833727$	$L_{(2)} = 0,738532$
	Критичні значення	$C_{5\%} = 0,193$	Приймається	Приймається

- 95 % – $1468,83 \pm 198,736$;
- 99 % – $1468,83 \pm 261,1833$.

За вищенаведеною методикою виконано оцифрування кількісних і якісних ознак ціноутворювальних чинників.

Вихідні дані для побудови оцінювальних моделей виділеного кластеру (22 об'єкти земельних ділянок Львівського регіону для садибного використання) подано в табл. 3.

На рис. 2 також представлено графік питомих вартостей виділеного кластеру (варіаційний ряд).

Візуальний аналіз даних не показує серед значень варіаційного ряду аномальних, але все таки необхідно виконати перевірку вибірки (табл. 3) статистичними методами та за допомогою нормального розподілу.

За допомогою критеріїв Смірнова–Грбса, Грбса та Тіт'сна–Мура виконаємо перевірку вибірки на грубі похибки (табл. 4).

Усі отримані у табл. 4 критерії свідчать, що мінімальне і максимальне значення вважаються для вибірки типовими з великим запасом. Виконані дослідження в табл. 4 дають змогу перевірити гіпотезу про нормальність розподілу даних. Розглянемо їх:

1. Перевірка за коефіцієнтом варіації. Для нашої вибірки значення коефіцієнта варіації становить 31,58 % (0,315812), тобто не перевищує значення в 33 %, що дає змогу виконувати перевірку за іншими критеріями.

2. Перевірка за критерієм середнього абсолютного відхилення (САВ). Для вибірки із 21 значення отримасмо $0,06025 < 0,087287$, що справедливо і гіпотеза про нормальність розподілу приймається.

3. Перевірка за розмахом варіацій. Отримане критеріальне відношення становить 2,82164. У разі зіставлення з табличними даними встановлено, що умова виконується на рівні значущості 10 % і знаходиться в межах інтервалу $[2,59; 3,31]$ і тому за критерієм розмаху варіацій гіпотеза про нормальність розподілу цілком прийнятна.

4. Перевірка за допомогою показників асиметрії та ексцесу. Для нашого випадку $|G_1| = 0,4345$; $S_{G_1} = 0,51197$ і $|G_2| = -1,3421$;

$S_{G_2} = 0,9720$, встановлено що показники асиметрії та ексцесу мають незміщені оцінки та їх середньоквадратичні відхилення виконуються одночасно. Так, гіпотеза про нормальність розподілу, за такими критеріями, може бути прийнятою.

Отже, виконана перевірка варіаційного ряду із 21 об'єкта не виявила аномальних значень питомих вартостей і гіпотеза про нормальність розподілу вибірових даних підтверджується. Це дає змогу перейти до побудови оцінювальних моделей методом регресивного аналізу.

Побудова оцінювальних моделей та перевірка їхньої якості

Метою дослідження є побудова економіко-математичних моделей методами регресивного аналізу, що дасть змогу встановити не тільки найкращу оцінювальну модель для цієї конкретної вибірки, але й встановити принципи моделювання вартості об'єктів нерухомості на основі відомих цін аналогів.

Реалізацію алгоритму регресивного аналізу необхідно розпочати з побудови матриці вибірових парних коефіцієнтів кореляції (табл. 5). Обчислення та аналіз коефіцієнтів кореляції дасть змогу встановити зв'язок вартості з кожним із чинників. Яка існує кореляція між значеннями чинників?

Як видно із табл. 5 питома вартість має досить сильний негативний зв'язок із відстанню до межі обласного центру (м. Львів) та відстані до межі районного центру; середню позитивну – з газифікацією і каналізацією. Чинник водопостачання не буде розглядатися тому, що для всіх об'єктів цього кластера його немає.

Незначущість зв'язку між питомою вартістю і такими важливими характеристиками, як якість електрифікації та відстань до залізничної станції говорить про те, що вплив цих чинників враховано під час формування цього кластеру. Серед незалежних змінних немає високорельєваних, що є позитивним чинником.

Побудова адитивних моделей

Формулу для побудови адитивної моделі, для нашого дослідження, можна записати у такому загальному вигляді:

$$\begin{aligned} C = & a_0 + a_1 \cdot S_{LVIV} + a_2 \cdot S_V + a_3 \cdot S_L + \\ & + a_4 \cdot G + a_5 \cdot E + a_6 \cdot K + a_7 \cdot S_R + a_8 \cdot S_S + \\ & + a_9 \cdot P, \end{aligned} \quad (9)$$

де C – вартість об'єкта оцінювання; $a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9$ – коефіцієнти моделі, які пов'язані із ціноутворювальними чинниками; $S_{LVIV}, S_V, S_L, G, E, K, S_R, S_S, P$ – виділені нами вище ціноутворювальні чинники.

Розрахунок дев'ятифакторної адитивної моделі подано в табл. 6.

Аналогічно на розрахували інші види адитивних моделей від восьмифакторної до однофакторної. Представимо ці моделі та виконаємо їх аналіз:

1. Дев'ятифакторна адитивна модель має такий вид.

$$\begin{aligned} C = & 2506,19 - 44,58 \cdot S_{LVIV} - 201,40 \cdot S_V + \\ & + 59,56 \cdot S_L + 271,98 \cdot G - 9,49 \cdot E + \\ & + 112,74 \cdot K - 31,98 \cdot S_R + 3,24 \cdot S_S - 1,13 \cdot P. \end{aligned}$$

Як видно із табл. 6 коефіцієнт визначеності моделі $R^2 = 0,9677$ стверджує, що побудована модель пояснює 96,77 % дисперсії цін аналогів. Це вважається надзвичайно добрим результатом для будь-якої моделі, особливо враховуючи малий об'єм вибірки ($n = 21$).

Значення коефіцієнта Фішера $F = 36,57$ стверджує, що регресивна залежність значна зі 95 % статистичною надійністю, причому із великим запасом, що говорить про існування тісного зв'язку між ціною та ціноутворювальними чинниками. Для виконання перевірки значущості окремих регресивних коефіцієнтів на основі t -критерію Стюдента необхідно визначити, за допомогою відповідних таблиць (таблиці розподілу Стюдента), його критичне значення $t_{кр}$.

Виконані дослідження показали, що тільки для трьох коефіцієнтів регресії за змінних (відстань до межі Львова, відстань до водойми і наявність та якість газифікації), $|t| > t_{кр}$. Для інших шести чинник значущості регресивних коефіцієнтів зі статистичною надійністю 95 % не підтверджується.

Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт за змінної S_S (відстань до залізничної станції) і для побудови восьмифакторної адитивної моделі його необхідно вилучити.

1. Восьмифакторна адитивна модель.

$$\begin{aligned} C = & 2510,91 - 44,24 \cdot S_{LVIV} - 200,26 \cdot S_V + \\ & + 60,47 \cdot S_L + 270,41 \cdot G - 12,78 \cdot E + \\ & + 112,84 \cdot K - 32,62 \cdot S_R - 1,37 \cdot P. \end{aligned}$$

Серед незначущих коефіцієнтів регресії найменш суттєвим за t -критерієм є коефіцієнт за змінної E (наявність і якість електрофікації) і для побудови семифакторної адитивної моделі його необхідно вилучити.

2. Семифакторна адитивна модель.

$$\begin{aligned} C = & 2506,34 - 44,96 \cdot S_{LVIV} - 195,56 \cdot S_V + \\ & + 62,85 \cdot S_L + 270,62 \cdot G + 117,09 \cdot K - \\ & - 31,17 \cdot S_R - 1,47 \cdot P. \end{aligned}$$

Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт за змінної P (площа об'єкта оцінки в сотках) і для побудови шестифакторної адитивної моделі його необхідно вилучити.

3. Шестифакторна адитивна модель.

$$\begin{aligned} C = & 2485,72 - 44,34 \cdot S_{LVIV} - 200,24 \cdot S_V + \\ & + 60,23 \cdot S_L + 260,74 \cdot G + 120,75 \cdot K - 31,96 \cdot S_R. \end{aligned}$$

Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт за змінної S_L (відстань до лісу) і для побудови п'ятифакторної адитивної моделі його необхідно вилучити.

4. П'ятифакторна адитивна модель.

$$\begin{aligned} C = & 2583,85 - 45,42 \cdot S_{LVIV} - 207,02 \cdot S_V + \\ & + 252,12 \cdot G + 115,78 \cdot K - 34,29 \cdot S_R. \end{aligned}$$

Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт за змінної K (забезпеченість каналізацією) і для побудови чотирифакторної адитивної моделі його необхідно вилучити.

5. Чотирифакторна адитивна модель.

$$\begin{aligned} C = & 2648,43 - 46,97 \cdot S_{LVIV} - 250,16 \cdot S_V + \\ & + 293,94 \cdot G - 35,65 \cdot S_R. \end{aligned}$$

Отриману модель вважатимемо базовою (основною) тому, що всі коефіцієнти регресії значущі. Для експерименту відсіювати чинники продовжимо доти, доки не буде отримана регресія, що враховує вплив одного ціноутворювального чинника. Найменш суттєвим зі всіх за значенням t -критерію є коефіцієнт за змінної S_V (відстань до водойми) для побудови трифакторної адитивної моделі його необхідно вилучити.

Таблиця 5

Матриця коефіцієнтів кореляції r

Table 5

The matrix of of the correlation coefficients

Чинники	Відстань до межі Львова, км	Відстань до водойми, км	Відстань до лісу, км	Газифікація	Електрифікація	Каналізація	Відстань до межі районного центру, км	Відстань до з/д станції, км	Розмір ділянки, сотки	Ціна за сотку, у.о.
Відстань до межі Львова, км	1	0,08131	-0,29236	-0,12692	0,52871	-0,3283	0,36515	0,32715	-0,42204	-0,85667
Відстань до водойми, км	0,081312	1	0,01049	0,26533	0,26426	-0,22502	-0,48244	0,33003	0,21775	-0,05779
Відстань до лісу, км	-0,29236	0,01049	1	0,02485	-0,22318	0,048141	-0,22368	0,14368	0,25179	0,33337
Газифікація	-0,12692	0,26533	0,02485	1	0,27735	0,395285	-0,61262	-0,09156	0,40635	0,51944
Електрифікація	0,52871	0,26426	-0,22318	0,27735	1	-0,25215	-0,37452	0,17267	-0,13697	-0,26224
Каналізація	-0,3283	-0,22502	0,048141	0,395285	-0,25215	1	-0,25239	-0,03185	0,06128	0,54797
Відстань до межі районного центру, км	0,36515	-0,48244	-0,22368	-0,61262	-0,37452	-0,25239	1	-0,23143	-0,26835	-0,63345
Відстань до з/д станції, км	0,32715	0,33003	0,14368	-0,09156	0,17267	-0,03185	-0,23143	1	-0,473	-0,21891
Розмір ділянки, сотки	-0,42204	0,21775	0,25179	0,40635	-0,13697	0,06128	-0,26835	-0,473	1	0,43270
Ціна за сотку, у.о.	-0,85667	-0,05779	0,33337	0,51944	-0,26224	0,54797	-0,63345	-0,21891	0,43270	1

Таблиця 6

Розрахунок дев'ятифакторної адитивної моделі

Table 6

Calculation of nine factorial additive model

Регресивна статистика	
Множинний R	0,983697
R-квадрат	0,967659
Нормований R-квадрат	0,941199
Стандартна помилка	112,6762
Спостереження	21

Дисперсійний аналіз	df	SS	MS	F	Значущість F
Регресія	9	4178577	464286,4	36,56972577	6,26E-07
Залишок	11	139655,1	12695,92		
Разом	20	4318232			

	Коефіцієнти	Стандартна похибка	f – статистика	P – значення	Нижні 95%	Верхні 95%	Нижні 95,0%	Верхні 95,0%
a_0	2506,191	183,8784	13,62962	3,11063E-08	2101,478	2910,905	2101,478	2910,905
Змінна X 1	-44,5807	9,130032	-4,88286	0,000484666	-64,6758	-24,4856	-64,6758	-24,4856
Змінна X 2	-201,406	88,79399	-2,26823	0,044445162	-396,84	-5,97127	-396,84	-5,97127
Змінна X 3	59,56168	54,77987	1,087291	0,300158535	-61,008	180,1314	-61,008	180,1314
Змінна X 4	271,976	85,67925	3,174351	0,008849754	83,39728	460,5548	83,39728	460,5548
Змінна X 5	-9,49515	114,6654	-0,08281	0,93549234	-261,872	242,8818	-261,872	242,8818
Змінна X 6	112,7364	80,534	1,399861	0,18912096	-64,5177	289,9906	-64,5177	289,9906
Змінна X 7	-31,9678	18,43968	-1,73364	0,1108811	-72,5533	8,617638	-72,5533	8,617638
Змінна X 8	3,24442	45,57811	0,071184	0,944529248	-97,0723	103,5612	-97,0723	103,5612
Змінна X 9	-1,12732	7,042231	-0,16008	0,875720182	-16,6272	14,37253	-16,6272	14,37253

Таблиця 7

Аналіз адитивних моделей

Table 7

Analysis of the additive models

Характеристика моделей		Кількість чинників в моделі								
Коефіцієнти регресії	Чинники	9	8	7	6	5	4	3	2	1
	a_0	2506,19	2510,91	2506,34	2485,73	2583,85	2648,43	2508,28	2415,21	2762,94
	Відстань до межі Львова, км	-44,58	-44,25	-44,96	-44,34	-45,42	-46,97	-51,84	-54,95	-58,57
	Відстань до водойми, км	-201,41	-200,26	-195,56	-200,24	-207,02	-250,16	0	0	0
	Відстань до лісу, км	59,56	60,47	62,85	60,25	0	0	0	0	0
	Газифікація	271,98	270,41	270,63	260,74	252,12	293,94	315,29	401,55	0
	Електрифікація	-9,50	-12,79	0	0	0	0	0	0	0
	Каналізація	112,74	112,84	117,09	120,75	115,78	0	0	0	0
	Відстань до межі районного центру, км	-31,97	-32,62	-31,17	-31,96	-34,29	-35,65	-18,10	0	0
	Відстань до з/д станції, км	3,24	0	0	0	0	0	0	0	0
	Розмір ділянки, сотки	-1,13	-1,37	-1,47	0	0	0	0	0	0
Статистичні властивості	R^2	96,77	96,76	96,76	96,74	96,28	96,50	91,84	90,53	73,39
	Стандартне відхилення	112,68	107,90	103,74	100,23	103,39	110,18	143,92	150,71	245,93
	F	36,57	44,86	55,46	69,31	77,79	84,92	63,82	86,06	52,40
Якість оцінки	Похибка апроксимації, %	-5,75	-6,03	-5,09	-5,99	-5,60	-2,86	4,03	14,52	6,10

6. Трифакторна адитивна модель.

$$Ц = 2508,28 - 51,84 \cdot S_{LVIV} + 315,29 \cdot G - 18,10 \cdot S_R.$$

Найменш суттєвим зі всіх за значенням t -критерію є коефіцієнт за змінної S_R (відстань до районного центру) для побудови двофакторної адитивної моделі його необхідно вилучити.

7. Двофакторна адитивна модель.

$$Ц = 2415,21 - 54,95 \cdot S_{LVIV} + 401,55 \cdot G.$$

Найменш суттєвим зі всіх за значенням t -критерію є коефіцієнт за змінної G (наявність та якість газопостачання) для побудови однофакторної адитивної моделі його необхідно вилучити.

8. Однофакторна адитивна модель.

$$Ц = 2762,94 - 58,57 \cdot S_{LVIV}.$$

Значення коефіцієнтів усіх побудованих регресій (адитивні моделі), їхні статистичні властивості і прогнози особливості ПОДАНО в табл. 7.

Побудова мультиплікативних моделей

Загальну формулу для побудови мультиплікативної моделі для нашого дослідження доцільно подати у такому загальному вигляді:

$$\ln(Ц) = \ln(a_0) + a_1 \cdot \ln(S_{LVIV}) + a_2 \times \ln(S_V) + a_3 \cdot \ln(S_L) + a_4 \cdot G + a_5 \cdot E + a_6 \times (10) \times K + a_7 \cdot \ln(S_R) + a_8 \cdot \ln(S_S) + a_9 \cdot \ln(P),$$

де $Ц$ – вартість об'єкта оцінки; $a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9$ – коефіцієнти моделі, які пов'язані із ціноутворювальними чинниками;

$S_{LVIV}, S_V, S_L, G, E, K, S_R, S_S, P$ – виділені нами вище ціноутворювальні чинники.

Виконавши логарифмування обох частин рівняння (10) отримаємо:

$$C = a_0 \cdot S_{LVIV}^{a_1} \cdot S_V^{a_2} \cdot S_L^{a_3} \times \\ \times a_4^G \cdot a_5^E \cdot a_6^K \cdot S_R^{a_7} \cdot S_S^{a_8} \cdot P^{a_9} \quad (11)$$

Розрахунок дев'ятифакторної мультиплікативної моделі подано в табл. 8.

Аналогічно ми розраховали інші види мультиплікативних моделей від восьми-факторної до однофакторної.

Подамо ці моделі та виконаємо їх аналіз:

1. Дев'ятифакторна мультиплікативна модель.

$$C = 24161,6 \cdot S_{LVIV}^{-1,0844} \cdot S_V^{-0,017} \cdot S_L^{0,0091} \times \\ \times 0,1177^G \cdot 0,2568^E \cdot 0,1542^K \times \\ \times S_R^{0,0655} \cdot S_S^{0,0216} \cdot P^{0,0338}$$

Гіпотеза про нормальність розподілу залишків з великим запасом підтверджується із довірчою імовірністю 0,95. Як видно із виконаних досліджень (табл. 5, табл. 8), як у адитивної, так і у мультиплікативної дев'ятифакторних моделях коефіцієнти визначеності ($R^2 = 0,9677$) і коефіцієнти Фішера ($F = 36$) повністю збігаються, що говорить про високу кореляцію досліджень.

За результатами перевірки значущості коефіцієнтів статистично істотними виявились тільки три змінні, а саме: відстань до межі обласного центру; наявність і якість електрифікації, наявність і якість каналізації.

Як видно із табл. 8 серед незначущих коефіцієнтів регресії найменш суттєвими за значенням t -критерію є коефіцієнт регресії за змінної S_L – відстань до лісу і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

2. Восьмифакторна мультиплікативна модель.

$$C = 24849,2 \cdot S_{LVIV}^{-1,0968} \cdot S_V^{-0,015} \times \\ \times 0,1129^G \cdot 0,2577^E \cdot 0,1557^K \times \\ \times S_R^{0,0665} \cdot S_S^{0,0239} \cdot P^{0,0374}$$

Для цієї моделі статистично суттєвими виявились вже чотири ціноутворювальні чинники, а саме: відстань до межі обласного центру; наявність і якість електрифікації, наявність і якість газопостачання, наявність і якість каналізації. Серед незначущих коефіцієнтів

регресії найменш суттєвими за значенням t -критерію є коефіцієнт регресії за змінної S_V – відстань до водойми і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

3. Семифакторна мультиплікативна модель.

$$C = 25675,4 \cdot S_{LVIV}^{-1,1091} \cdot 0,1098^G \cdot 0,2640^E \times \\ \times 0,1640^K \cdot S_R^{0,0776} \cdot S_S^{0,0234} \cdot P^{0,0368}$$

Для цієї моделі статистично суттєвими виявились ті ж самі чотири ціноутворювальні чинники, що і для попередньої моделі. Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт регресії за змінної P – площа об'єкта оцінки і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

4. Шестифакторна мультиплікативна модель.

$$C = 28570,8 \cdot S_{LVIV}^{-1,1144} \cdot 0,1234^G \times \\ \times 0,2577^E \cdot 0,1582^K \cdot S_R^{0,0773} \cdot S_S^{0,0195}$$

Для цієї моделі статистично суттєвими виявились ті самі чотири ціноутворювальні чинники, що і для двох попередніх моделей. Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт регресії за змінної S_S – відстань до залізничної станції і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

5. П'ятифакторна мультиплікативна модель.

$$C = 25190,9 \cdot S_{LVIV}^{-1,0677} \cdot 0,1137^G \times \\ \times 0,2529^E \cdot 0,1631^K \cdot S_R^{0,0635}$$

Для цієї моделі статистично суттєвими виявились ті самі чотири ціноутворювальні чинники, що і для трьох попередніх моделей. Серед незначущих коефіцієнтів регресії найменш суттєвим за значенням t -критерію є коефіцієнт регресії за змінної S_S – відстань до районного центру і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

6. Чотирифакторна мультиплікативна модель.

$$C = 19764,9 \cdot S_{LVIV}^{-0,9402} \times \\ \times 0,0919^G \cdot 0,1924^E \cdot 0,1649^K$$

Серед незначущих коефіцієнтів регресії найменш суттєвими за значенням t -критерію є коефіцієнт регресії за змінної G – наявність і якість газопостачання і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

Таблиця 8

Table 8

Розрахунок дев'ятифакторної мультиплікативної моделі

Calculation nine factorial multiplicative model

Регресивна статистика	
Множинний R	0,983723
R-квадрат	0,967711
Нормований R-квадрат	0,941293
Стандартна помилка	0,076614
Спостереження	21

Дисперсійний аналіз	df	SS	MS	F	Значущість F
Регресія	9	1,935098	0,215011	36,63049	6,2E-07
Залишок	11	0,064567	0,00587		
Разом	20	1,999665			

Змінна	Коефіцієнти	Стандартна похибка	f – статистика	P – значення	Верхні 95%		Верхні 95,0 %	
					Нижні 95 %	Верхні 9 5%	Нижні 95,0 %	Верхні 95,0 %
a_0	10,09252	0,342983	29,42572	8,23E-12	9,337617	10,84742	9,337617	10,84742
Змінна X1	-1,0844	0,118358	-9,16204	1,76E-06	-1,3449	-0,8239	-1,3449	-0,8239
Змінна X 2	-0,01698	0,027329	-0,62147	0,546953	-0,07713	0,043166	-0,07713	0,043166
Змінна X 3	0,009181	0,023794	0,385837	0,706975	-0,04319	0,06155	-0,04319	0,06155
Змінна X 4	0,117658	0,054798	2,147125	0,054921	-0,00295	0,238267	-0,00295	0,238267
Змінна X 5	0,256821	0,062562	4,105034	0,001744	0,119122	0,39452	0,119122	0,39452
Змінна X 6	0,154226	0,05056	3,050354	0,011041	0,042944	0,265508	0,042944	0,265508
Змінна X 7	0,065467	0,044427	1,473591	0,168624	-0,03232	0,163249	-0,03232	0,163249
Змінна X 8	0,021634	0,01826	1,18477	0,261075	-0,01856	0,061824	-0,01856	0,061824
Змінна X 9	0,033794	0,061051	0,553543	0,590969	-0,10058	0,168166	-0,10058	0,168166

7. Трифакторна мультиплікативна модель.

$$Ц = 23604,2 \cdot S_{LVIV}^{-0,9921} \cdot 0,2444^E \cdot 0,2067^K.$$

Отримана модель вважається нами базовою (основною) тому, що всі коефіцієнти регресії значущі. Для експерименту процес відсіву чинників продовжимо доти, доки не буде отримана регресія, що враховує вплив одного ціноутворювального чинника. Серед незначущих коефіцієнтів регресії найменш суттєвими за значенням t -критерію є коефіцієнт регресії за змінної K – наявність і якість каналізації і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

8. Двофакторна мультиплікативна модель.

$$Ц = 14967,3 \cdot S_{LVIV}^{-0,7916} \cdot 0,1878^E.$$

Серед незначущих коефіцієнтів регресії найменш суттєвими за значенням t -критерію є коефіцієнт регресії за змінної E – наявність і якість електрофікації і тому цей ціноутворювальний чинник необхідно вилучити із наступної моделі.

9. Однофакторна мультиплікативна модель.

$$Ц = 20712,6 \cdot S_{LVIV}^{-0,8835}.$$

Значення коефіцієнтів усіх побудованих регресій (мультиплікативні моделі), їхні статистичні властивості і прогностні особливості подано в табл. 9.

Як видно із табл. 7 і 9 для всіх побудованих моделей, як адитивних, так і мультиплікативних значення F – статистики значно перевищують табличні, що говорить про існування значного кореляційного зв'язку між ціною та ціноутворювальними чинниками. Значення цієї статистики із зменшенням кількості чинників з початку зростає, а потім спадає та досягає максимуму в моделях, які мають тільки значущі чинники. Коефіцієнт визначеності моделі R^2 із відсівом незначущих чинників незначно погіршується. Значення стандартного відхилення із зменшенням кількості чинників з початку зменшується, а пізніше зростає.

Таблиця 9

Аналіз мультиплікативних моделей

Table 9

Analysis of the multiplicative models

Характеристика моделей		Кількість чинників у моделі								
Чинники		9	8	7	6	5	4	3	2	1
Коефіцієнти регресії	a_0	24161,6	24849,2	25675,4	28570,8	25190,9	19764,9	23604,2	14967,3	20712,6
	Відстань до межі Львова, км	-1,08	-1,10	-1,11	-1,11	-1,07	-0,94	-0,99	-0,79	-0,88
	Відстань до водойми, км	-0,02	-0,01	0	0	0	0	0	0	0
	Відстань до лісу, км	0,01	0	0	0	0	0	0	0	0
	Газифікація	0,12	0,11	0,11	0,12	0,11	0,09	0	0	0
	Електрифікація	0,26	0,26	0,26	0,26	0,25	0,14	0,24	0	0
	Каналізація	0,15	0,16	0,16	0,16	0,16	0,16	0,20	0,19	0
	Відстань до межі районного центру, км	0,07	0,07	0,078	0,077	0,06	0	0	0	0
	Відстань до з/д станції, км	0,02	0,02	0,02	0,02	0	0	0	0	0
	Розмір ділянки, сотки	0,03	0,04	0,04	0	0	0	0	0	0
Статистичні властивості	R^2	96,77	96,72	96,64	96,53	96,09	97,60	94,08	83,73	72,85
	Стандартне відхилення	0,08	0,074	0,07	0,07	0,07	0,08	0,08	0,13	0,158
	F	36,63	44,34	53,37	64,88	73,76	80,48	90,00	46,32	66,79
Якість оцінки	Похибка апроксимації, %	13,76	13,70	15,14	16,60	16,13	4,74	2,87	7,24	2,12

Отже, найкращі прогнознi моделi – це моделi, що отриманi пiсля видiлення всiх незначущих чинникiв, тобто: чотирифакторна адитивна модель та трифакторна мультиплiкативна модель. Щодо сукупностi критерiїв цi моделi доцiльно вважати найкращими для опису цього кластеру об'єктiв оцiнки.

Пiсля вибору найкращих моделей iх необхідно перебудувати для можливостi встановлення вартостi будь-якого об'єкта оцiнки.

Чотирифакторна адитивна модель, що перебудована для цiлого кластера (22 спостереження) має вигляд:

$$C = 2639,97 - 47,3395 \cdot S_{LVIV} - 245,715 \cdot S_V + 302,6783 \cdot G - 33,9778 \cdot S_R; \\ (R^2 = 95,48, F = 89,69, CB = 0,0721).$$

Трифакторна мультиплiкативна модель, що перебудована для цiлого кластера (22 спостереження) має вигляд:

$$C = 23316,46 S_{LVIV}^{-0,98802} \cdot 0,2414^E \cdot 0,2085^K; \\ (R^2 = 94,05, F = 94,91, CB = 0,0814).$$

Коефiцiєнт при S_{LVIV} дорiвнює – 47,34, означає, що збiльшення вiдстанi вiд межi м. Львова на 1 %, зменшує цiну на 0,47 %, тобто еластичнiсть цiни об'єкта оцiнки iз вiддалення вiд межi Львова становить – 0,47. Для цiноутворювальних чинникiв S_V вiдстань до водного об'єкта та S_R вiдстань до районного центру будуть аналогiчнi до висновкiв (збiльшення вiдстанi вiд об'єкта оцiнювання до водного об'єкта або районного центра веде до зменшення вартостi цього об'єкта). Позитивний коефiцiєнт при G наявнiсть газопостачання, доводить, що цiна земельної дiлянки зростає за можливостi пiдключення до магiстрального газопроводу.

Аналіз розміру і структури похибок оціночних моделей подано в табл. 10.

На рис. 5 подано результати економіко-математичного моделювання вибірки із 22 об'єктів оцінки (вибраний кластер), що відповідають сегменту ринку нерухомості Львівської області – землі для садибної забудови.

Таблиця 10

Аналіз розміру і структури похибок для адитивної та мультиплікативної моделей

Table 10

Analysis of the size and structure of errors for additive and multiplicative models

№ з/п	Фактичні значення цін об'єктів, у.о.	Розраховані адитивною моделлю, у.о.	Різниця, у.о.	Різниця, %	Розраховані мультиплік. моделлю, у.о.	Різниця, у.о.	Різниця, %
1	2143	2115	28	1,31	2044	99	4,62
2	2080	2051	29	1,39	2002	78	3,75
3	2200	2144	56	2,55	2466	-266	-12,09
4	2167	2129	38	1,75	2117	50	2,31
5	2133	2048	85	3,98	1978	155	7,27
6	1667	1665	2	0,12	1538	129	7,74
7	1400	1323	77	5,50	1400	0	0,00
8	1467	1482	-15	-1,02	1520	-53	-3,61
9	909	848	61	6,71	885	24	2,64
10	917	815	102	11,12	938	-21	-2,29
11	1600	1482	118	7,38	1520	80	5,00
12	1625	1625	0	0,00	1400	225	13,85
13	1250	1362	-112	-8,96	1538	-288	-23,04
14	1792	1956	-164	-9,15	1806	-14	-0,78
15	889	849	40	4,50	885	4	0,45
16	1000	1100	-100	-10,00	1100	-100	-10,00
17	1083	1125	-42	-3,88	1100	-17	-1,57
18	1100	1125	-25	-2,27	1100	0,00	0,00
19	1200	1100	100	8,33	1100	100	8,33
20	1067	1358	-291	-27,27	1103	-36	-3,37
21	1158	1173	-15	-1,30	1187	-29	-2,50
22	1500	1473	27	1,80	1538	-38	-2,53

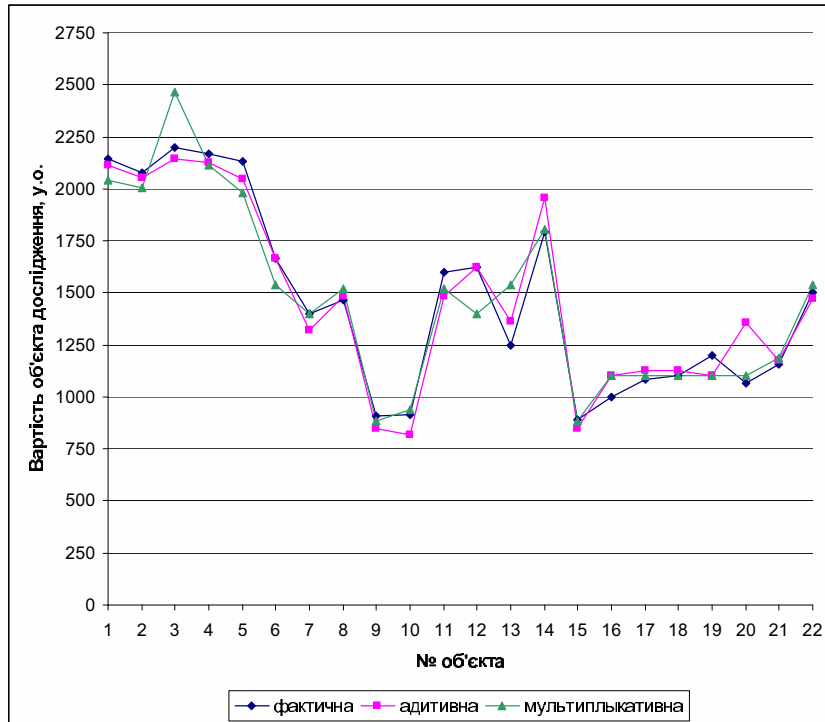


Рис. 5. Результат економіко-математичного моделювання

Fig. 5. The result of economic and mathematical modeling

Таблиця 11

Похибки моделей для адитивної та мультиплікативної моделей

Table 11

Errors of models for additive and multiplicative models

№ з/п	Похибки адитивної моделі Δ_A	Δ_A^2	Похибки мультиплікативної моделі Δ_M	Δ_M^2
1	28	784	99	9801
2	29	841	78	6084
3	56	3136	-266	70756
4	38	1444	50	2500
5	85	7225	155	24025
6	2	4	129	16641
7	77	5929	0	0
8	-15	225	-53	2809
9	61	3721	24	576
10	102	10404	-21	441
11	118	13924	80	6400
12	0	0	225	50625
13	-112	12544	-288	82944
14	-164	26896	-14	196
15	40	1600	4	16
16	-100	10000	-100	10000
17	-42	1764	-17	289
18	-25	625	0,00	0
19	100	10000	100	10000
20	-291	84681	-36	1296
21	-15	225	-29	841
22	27	729	-38	1444
	$[\Delta_A^2 = 196701]$		$[\Delta_M^2 = 297684]$	

Отже, ми отримали кінцеві оцінювальні моделі для оцінювання всіх можливих неоцінених об'єктів цього кластеру. Для того, щоб мати можливість передбачення можливих похибок моделей, необхідно перевірити структуру і розмір похибок, отриманих як різниці між фактичними та розрахунковими значеннями.

Як видно із рис. 5 та табл. 11 для моделей виявилось по три похибки з відхиленням більше ніж 10 % від фактичних значень вартості. Отримані моделі мають похибки під час оцінки різних об'єктів, тобто якщо одна модель має похибку, то інша модель для того самого об'єкта дає хороший результат. Для практичних розрахунків доречно, на нашу думку, застосовувати обидві моделі.

Середня квадратична похибка є найкращим критерієм для оцінювання точності варіаційного ряду. Вона має такі переваги над середньою та імовірними похибками:

- середня квадратична похибка є досить чутливою мірою точності тому, що на її величину суттєвіше впливають великі за абсолютною величиною випадкові похибки, що переважно і визначають надійність результатів вимірювань;
- середня квадратична похибка вже за деякої порівняно не дуже великої кількості значень набуває сталого значення і в подальшому за збільшення кількості значень змінюється дуже повільно;
- за величиною середньої квадратичної похибки можна знайти граничну похибку, тобто таке найбільше за абсолютною величиною значення випадкової похибки, яке ще може з'явитися за цих умов і за граничну приймається потрібна середня квадратична похибка.

Розмір похибок оціночних моделей з використанням кластеризації подано на рис. 6.

Виконаємо дослідження отриманих похибок двома моделями на випадковість.

1. Обчислимо середню квадратичну похибку. Середньою квадратичною похибкою m називається величина, яка дорівнює квадратичному кореню зі середнього арифметичного квадратів істинних похибок.

$$m = \pm \sqrt{\frac{[\Delta^2]}{n}}, \quad (10)$$

де Δ – істинні похибки; n – кількість спостережень.

Отже, середня квадратична похибка для адитивної моделі становить $m_A = \pm 94 \text{ у.о.}$, а для мультиплікативної моделі – $m_M = \pm 116 \text{ у.о.}$

2. Знайдемо середню похибку за формулою $\theta = \frac{[\Delta]}{n}$. Для адитивної моделі ця похибка становить $\theta_A = 69 \text{ у.о.}$, а для мультиплікативної моделі – $\theta_M = 82 \text{ у.о.}$

3. Розрахуємо ймовірну похибку за формулою $\rho = \frac{|x_{e1}| + |x_{e2}|}{2}$, де x_{e1} , x_{e2} – значення похибок, що знаходяться в середині варіаційного ряду. Отже, для адитивної моделі ймовірна похибка становитиме $\rho_A = 41 \text{ у.о.}$, а для мультиплікативної моделі – $\rho_M = 44 \text{ у.о.}$

4. Перевіримо виконання умов $\theta \approx \frac{4}{5} \cdot m$ і $\rho \approx \frac{2}{3} \cdot m$, тобто

$$\frac{\theta_A}{m_A} = \frac{69}{94} = 0,73, \quad \frac{\theta_M}{m_M} = \frac{82}{116} = 0,71,$$

$$\frac{\rho_A}{m_A} = \frac{41}{94} = 0,44, \quad \frac{\rho_M}{m_M} = \frac{44}{116} = 0,38.$$

5. Обчислимо граничну похибку $\Delta_{\text{гран}} = 3 \cdot m$. Отже, для адитивної моделі гранична похибка становить $\Delta_{\text{гран}} = 3 \cdot m_A = 3 \cdot (\pm 94) = \pm 282 \text{ у.о.}$, а для мультиплікативної моделі – $\Delta_{\text{гран}} = 3 \cdot m_M = 3 \cdot (\pm 116) = \pm 348 \text{ у.о.}$

Гранична похибка не перевищує (крім одного значення) похибки адитивної і мультиплікативної моделей.

Результати

Виконані дослідження доводять важливість застосування методики кластеризації для впровадження в Україні масової оцінки нерухомості. Якість економіко-математичних моделей визначають характеристики адекватності, стійкості та корисності моделі, які можна трактувати як погодження інформації, що описує функціональні можливості моделі, з наявною інформацією про реальний об'єкт дослідження та мету моделювання.

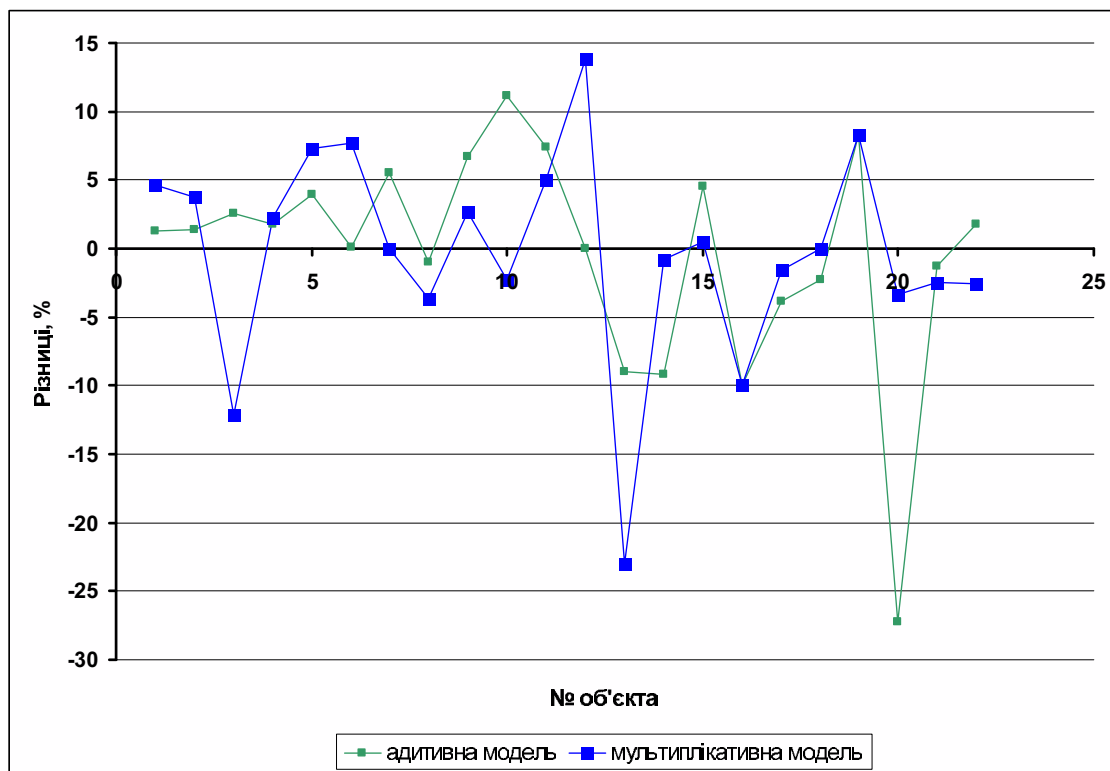


Рис. 6. Розмір похибок оцінювальних моделей з використанням кластеризації
 Fig. 6. The value of error of evaluation models using clustering

Наукова новизна

Виконані дослідження дають змогу отримувати кінцеві оцінювальні моделі для оцінювання неоцінених об'єктів будь-якого кластеру. Побудова моделі передбачає ретельне збирання ринкової інформації і детальне аналітичне опрацювання. Під час побудови моделі з метою отримання оптимального варіанта моделі, що дасть змогу уникнути недовраховання наявних чинників та не допустить включення в модель надто великої кількості змінних пропонується зменшити кількість чинників з використанням багатокрокового регресивного аналізу.

Практична значущість

У роботі виконано практичну реалізацію теоретичних розробок. Здійснено побудову оцінювальних моделей методами регресивного аналізу і перевірку якості оцінки на прикладі відібраного кластеру (садибна забудова) Львівського регіону. Подані в роботі дослідження доцільно застосовувати для населених пунктів України.

Висновки

1. Удосконалено методику кластеризації, як найперспективнішого методу регресивного аналізу. Порівняльний аналіз економіко-математичних моделей показав, що для розрахунку масової вартості нерухомості доцільно обмежитися використанням регресивних моделей, тому що незважаючи на їх відносну простоту, за якістю оцінки не є гіршими від моделей інших класів.

2. Побудовано комплекс оцінювальних моделей для розрахунку кадастрової вартості нерухомості (сегменту земель садибної забудови) Львівського регіону. Найкращими виявились моделі, отримані після усунення всіх незначущих ціноутворювальних чинників. Аргументовано доведено їх доцільність для розрахунку вартості не оцінених об'єктів нерухомості.

3. Виконавши побудову моделей для кожного кластера оцінюваної території, ми отримали можливість розраховувати вартість будь-якого об'єкта оцінки, шляхом підстановки параметрів об'єктів оцінки у рівняння регресії і розрахунку вартості за відповідними регресивними коефіцієнтами.

4. Якщо отримана вартість об'єкта різними методами приблизно однакова за кінцевий результат доцільно вважати середнє значення, якщо отримані результати суттєво відрізняються, який результат вважати більш коректним вирішувати експерту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Міжнародні стандарти оцінки МСО-1. Ринкова вартість як база оцінки. Міжнародний комітет зі стандартів оцінки майна (IVSC), 2006 р.
- Міжнародні стандарти оцінки МСО-2. Загальні поняття і принципи оцінки. – Міжнародний комітет зі стандартів оцінки майна (IVSC), 2006 р.
- Національний стандарт № 1 “Загальні засади оцінки майна і майнових прав”, затверджений постановою Кабінету Міністрів України від 10.09.2003 р. № 1440.
- Національний стандарт № 2 “Оцінка нерухомого майна”, затверджений постановою Кабінету Міністрів України від 28.10.2004 р. № 1442.
- Губар Ю. Розробка підходів і методів кадастрової оцінки нерухомості населених пунктів / Ю. Губар // Сучасні досягнення геодезичної науки та виробництва. – 2012. – № II (24) – 146–150.
- Перович Л. М. Оцінка нерухомості : навч. посіб. / Л. М. Перович, Ю. П. Губар. – Львів : Національний університет “Львівська політехніка”, 2010. – 296 с.
- Губар Ю. Вплив ринкових факторів на вартість земель в локальних районах міста Львова / Ю. Губар // Сучасні досягнення геодезичної науки та виробництва. – 2008. – № II(16). – С. 157–162.
- Харрисон Г. С. Оценка недвижимости: Учебное пособие / Харрисон Генри [Пер. с англ.] – М. : РИО Мособлупрполиграфиздата, 1994. – 231 с.
- Организация оценки и налогообложения недвижимости / под ред. Джозефа К. Эккерта. – М. : Российское общество оценщиков, Академия оценки, Стар Интер, 1997. – Т. 1. – 382 с. – Т. 2 – 442 с.
- Организация оценки недвижимости / под общ. ред. Дж. К. Эккерта. – М., РОО, 1999 – 325 с.
- Губар Ю. Дослідження кінцевих результатів коригування вартості об'єктів нерухомості в порівняльному підході / Ю. Губар // Сучасні досягнення геодезичної науки та виробництва. – 2013. – № II(26). – С. 117–121.
- Губар Ю. Визначення коефіцієнтів коригування за просторовими критеріями в порівняльному підході / Ю. Губар // Сучасні досягнення геодезичної науки та виробництва. – 2013. – № I(25). – С. 128–133.
- Губар Ю. Застосування багатовимірної моделі методу парних продаж для дослідження ринкової вартості об'єктів комерційної нерухомості на рівні обласного центру / Ю. Губар // Геодезія, картографія та аерофотознімання. Міжвідомчий наук.-техн. зб. – 2012. – № 76. – С. 133–136.
- Jack P. Friedman, Jack C. Harris, J. Bruce Lindeman. Dictionary of Real Estate Terms.–Third Edition. Barron's Educational Series, Inc., New York, 2012.
- Jack P. Friedman, Nicolas Ordway. Income Property Appraisal and Analysis. – American Society of Appraisers. – Prentice Hall, New Jersey, 2013.
- The Appraisal of Real Estate. Twelfth Edition – Chicago: Appraisal Institute, 2011. – 759 p.

Ю. ГУБАР

Кафедра кадастра территорий, Национальный университет “Львовская политехника”, ул. С. Бандеры, 12, Украина, 79013, тел. +38 (032) 2582631, эл. почта Yurii.P.Hubar@lpnu.ua

ВНЕДРЕНИЕ МЕТОДИКИ КЛАСТЕРИЗАЦИИ ДЛЯ ПОСТРОЕНИЯ ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ РЕГРЕССИОННЫХ МОДЕЛЕЙ ОЦЕНКИ НЕДВИЖИМОСТИ

Цель. Целью исследования является построение экономико-математических моделей методами регрессионного анализа, что позволит установить не только лучшую оценочную модель для данной конкретной выборки, но и установить принципы моделирования стоимости объектов недвижимости на основе известных цен аналогов. Реализацию алгоритма регрессионного анализа необходимо начать с построения матрицы выборочных парных коэффициентов корреляции. Исчисление и анализ коэффициентов корреляции позволят установить связь стоимости с каждым из факторов. **Методика.** Методика основывается на применении метода дендрограм с использованием пространственно сжатых факторов. С целью выбора метода группировки объектов недвижимости целесообразно провести сравнительный анализ и тестирование различных методов кластерного анализа. **Результаты.** Выполненные исследования доказывают важность применения методики кластеризации для внедрения в Украине массовой оценки недвижимости. Качество экономико-математических моделей определяют характеристики адекватности, устойчивости и полезности

модели, которые можно трактовать как согласование информации, описывающей функциональные возможности модели. **Научная новизна.** Выполненные исследования позволяют получать конечные оценочные модели для оценки неопцененных объектов любого кластера. Построение модели предусматривает тщательный сбор рыночной информации и детальное аналитическая обработка. В процессе построения модели с целью получения оптимального варианта модели предлагается уменьшить количество факторов с использованием многошагового регрессионного анализа. **Практическая значимость.** Осуществлено построение оценочных моделей методами регрессионного анализа и проверку качества оценки на примере Львовского региона. Исследования целесообразно применять для населенных пунктов Украины.

Ключевые слова: оценка недвижимости, экономико-математическое моделирование, методика кластеризации, регрессивные модели.

YU. HUBAR

Department of Cadastre areas, Lviv Polytechnic National University, S. Bandera str., 12, Lviv, Ukraine, 79013, tel. +38 (032) 2582631, email: Yurii.P.Hubar@lpnu.ua

CLUSTERING IMPLEMENTATION OF THE METHODOLOGY FOR THE CONSTRUCTION OF ECONOMIC AND MATHEMATICAL REGRESSION MODEL OF VALUATION REAL ESTATE

Purpose. The purpose of the research is to construct economic and mathematical models by regression analysis, which will determine not only the best model for valuation of this particular sample, but also to establish principles of modelling the value real estate prices on the basis of known peers. Realization of the algorithm of regression analysis must begin with the construction of the matrix of paired sample correlation coefficients. The calculation and analysis of the correlation coefficients will allow to communicate the value of each of the factors. **Methodology.** The method is based on the method of dendrogram using the space compression factors. For the purpose of select the method of grouping of real estate is expedient conduct a comparative analysis and testing various methods of of cluster analysis. **Results.** Executed researches demonstrate the importance of using clustering methods for the introduction in Ukraine the cadastral valuation of real estate. The quality of economic and mathematical models define the characteristics of the adequacy, sustainability and usefulness models, which can be interpreted as approval of the information that describes the functionality of the model. **Originality.** Completed studies allow to obtain the final evaluation model to assess the invaluable objects of any cluster. Construction of the model includes a thorough collection of market information and detailed analytical processing. During the constructing a model in order to obtain the optimal variant of the model is proposed to reduce the number of factors by using multi-stage regression analysis. **Practical significance.** Implemented construction of evaluation models by regression analysis and quality control evaluation on the example of Lviv region. The research are useful for towns of Ukraine in order to determine the cadastral value of real estate.

Keywords: real estate market extraction, economic and mathematical modelling, methods of clustering, regression model.

REFERENCES

- Mizhnarodni standarty otsinky MSO-1. Rynkova vartist' yak baza otsinky* [International Valuation Standards]. *Mizhnarodnyy komitet zi standartiv otsinky mayna (IVSC)*, 2006.
- Mizhnarodni standarty otsinky MSO-2. Zahal'ni ponyattya i pryntsypy otsinky.* [International Valuation Standards]. *Mizhnarodnyy komitet zi standartiv otsinky mayna (IVSC)*, 2006.
- Natsional'nyy standart #1 "Zahal'ni zasady otsinky mayna i maynovykh prav"* [National Valuation Standards]. *zatverdzhenny postanovoyu Kabinetu Ministriv Ukrayiny vid 10.09.2003*, no. 1440.
- Natsional'nyy standart #2 "Otsinka nerukhomoho mayna"* [National Valuation Standards]. *zatverdzhenny postanovoyu Kabinetu Ministriv Ukrayiny vid 28.10.2004*, no. 1442.
- Hubar Yu. *Rozrobka pidkhodiv i metodiv kadaastrovoyi otsinky nerukhomosti naselenykh punktiv* [Development of approaches and methods cadastral valuation of real estate settlements]. *Suchasni dosyahnennya heodezychnoyi nauky ta vyrobnytstva* [Journal: "Modern geodesic advances of science and industry"]. 2012, no. II (24), pp. 146–150.
- Perovych L. M., Hubar Yu. P. *Otsinka nerukhomosti* [Real estate evaluation]: *navchal'nyy posibny*. Lviv: Lviv Polytechnic Publishing House“ 2010, 296 p.

- Hubar Yu. *Vplyv rynkovykh faktoriv na vartist' zemel' v lokal'nykh rayonakh mista L'vova* [The impact of for market factors on the value of lands in local areas of the city of Lviv]. *Suchasni dosyahnennya heodezychnoyi nauky ta vyrobnytstva* [Journal: "Modern geodesic advances of science and industry"]. 2008, no.II(16), pp. 157–162.
- Kharryson H. S. *Otsenka nedvyzhymosty*: [Real estate evaluation] *Uchebnoe posobyе*. Moscow: RYO Mosobluprpol'yhrafyzdata, 1994, 231 p.
- Orhanyzatsyya otsenky y nalohooblozhenyya nedvyzhymosty* [Organization evaluation and taxation of real estate]. Pod redaktsyyey Dzhozefa K. Эkkerta. Moscow, Rossyyskoe obshchestvo otsenshchykov, Akademyya otsenky, Star Ynter, 1997, T. 1 – 382 p., T. 2. – 442 p.
- Orhanyzatsyya otsenky nedvyzhymosty* [The organization of real estate evaluation]. Pod obshch. redaktsyyey Dzh. K. Эkkerta. Moscow, ROO, 1999, 325 p.
- Hubar Yu. *Doslidzhennya kintseyvkh rezul'tativ koryhuvannya vartosti ob"yektiv nerukhomosti v porivnyal'nomu pidkhodi* [The study final results correction value of real estate in the comparative approach]. *Suchasni dosyahnennya heodezychnoyi nauky ta vyrobnytstva* [Journal: "Modern geodesic advances of science and industry"]. 2013, no. II(26), pp. 117–121.
- Hubar Yu. *Vyznachennya koefitsiyentiv koryhuvannya za prostоровymy kryteriyamy v porivnyal'nomu pidkhodi* [Determination of the correction coefficients for dimensional criteria in the comparative approach]. *Suchasni dosyahnennya heodezychnoyi nauky ta vyrobnytstva* [Journal: "Modern geodesic advances of science and industry"]. 2013, no. I(25), pp. 128–133.
- Hubar Yu. *Zastosuvannya bahatovymirnoyi modeli metodu parnykh prodazh dlya doslidzhennya rynkovoyi vartosti ob"yektiv komertsyynoyi nerukhomosti na rivni oblasnoho tsentru* [The use of a multidimensional model of the method of paired sales to research the market value of commercial real estate of the regional center]. *Heodeziya, kartohrafiya ta aerofotoznimannya* [Geodesy, Cartography and Aerial Photography]. *Mizhvidomchyy nauk.-tekhn. Zb.*, 2012, no. 76.
- Jack P. Friedman, Jack C. Harris, J. Bruce Lindeman. *Dictionary of Real Estate Terms*. Third Edition. Barron's Educational Series, Inc., New York, 2012.
- Jack P. Friedman, Nicolas Ordway. *Income Property Appraisal and Analysis*. American Society of Appraisers, Prentice Hall, New Jersey, 2013.
- The Appraisal of Real Estate*. Twelfth Edition. Chicago: Appraisal Institute, 2011.

Надійшла 21.09.2015 р.