УДК 004.832.32.

Кравець І. О., Лисенко А. Ю.

ДОСЛІДЖЕННЯ МЕТОДУ ГРУПОВОГО УРАХУВАННЯ АРГУМЕНТІВ ДЛЯ АВТОРЕГРЕСІЙНИХ ТА ДИСТРИБУТИВНО-ЛАГОВИХ МОДЕЛЕЙ

Запропоновано удосконалення методу групового урахування аргументів (чітка та нечітка модифікації) для авторегресійних та дистрибутивно-лагових моделей, проведено системне наукове дослідження ефективності методу групового урахування аргументів (МГУА) для цих моделей. Також вперше було запропоновано модифікацію нечітких алгоритмів МГУА, що полягає у використанні двоїстого симплекс-методу замість звичайного симплекс-методу, що дозволило скоротити нечіткий МГУА з нечіткими входами та отримати покращені результати.

Ключові слова: *метод групового врахування аргументів, авторегресійні и дистрибутивно-лагові моделі.*

Предложено применение метода групового учета аргументов (четкая и нечеткая модификации) для авторегрессионных и дистрибутивно-лаговых моделей. Проведено исследование эфективности метода групового учета аргументов (МГУА) для этих моделей. Предложена модификация нечетких алгоритмов МГУА, с применением двойственного симплекс-метода вместо обычного симплекс-метода, что позволило сократить нечеткий МГУА з нечеткими входами и получить более точные результати.

Ключевые слова: *метод групового учета аргументов, авторегрессионные и дистрибутивно-лаговые модели.*

The expansion of the group method of data handling (its ordinary and fuzzy modifications) for autoregressive and lag's models is proposed. A systematic scientific study of the efficacy of the group method of data handling (GMDH) for autoregressive models is done. There is first proposed modification of the algorithm is to use dual-simplex method instead of the usual simplex-method, which reduced fuzzy GMDH with fuzzy inputs and get superior results.

Key words: *group method of data handling, autoregressive and lag's models.*

Вступ

На сьогоднішній день розроблені методи статистичного прогнозування, які дозволяють з високою точністю прогнозувати практично всі можливі часові ряди. Але вони базуються на математичному апараті, якій може бути використаний тільки при достатньому обсязі статистичних даних. Часто обсяг даних є недостатнім для пошуку залежностей для великої кількості факторів. Тому великої популярності для конкретних завдань прогнозування набуває метод групового урахування аргументі. який заснований на принципах «селекції» та гібридизації [1]. Цей метод дозволяє отримати коефіцієнти моделі при недостатньому обсязі даних та має більшу точність прогнозу. Однак, для авторегресійних та дистрибутивно-лагових моделей алгоритми МГУА відсутні, хоч при статистичному аналізі та прогнозуванні соціально-еконономічних показників часто необхідно враховувати інерційність та гальмування факторів.

У регресійному аналізі, якщо регресійна модель включає не лише поточні, а й попередні (лагові, або затримані) значення незалежних змінних (*x*), вона має назву ***дистрибутивно-лагова модель(ДЛМ).*** У той же час, якщо до моделі включене одне або більше попередніх значень залежної змінної *(y),* вона має назву ***авторегресійна модель.*** Таким чином, ***дистрибутивно-лагова модель****:*

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \varepsilon_t; \tag{1}$$

авторегресійна моделль:

$$y_t = \alpha + \beta x_t + \gamma y_{t-1} + \varepsilon_t$$

де ε_t – випадкова величина.

Шляхом перетворень дистрибутивно-лагова модель може бути зведена до авторегресійної моделлі. Такі авторегресійні моделі також відомі під назвою ***динамічних моделей,*** оскільки вони відображають часові зміни залежної змінної щодо її попереднього (попередніх) значенью Але нажаль традиційні методи статистичного аналізу (метод найменших квадратів) дають великі похибки та подібних моделей внаслідок кореляції випадкових похибок від спостереження до спостереження. Тому аналіз та розробка методів аналізу та прогнозування, придатних для авторегресійних моделей та при недостатньому обсязі статистичних даних є актуальною задачею.

Задачі дослідження

Задачею дослідження є модифікація методу групового урахування аргументів та виявлення доцільності прогнозування авторегресійних та дистрибутивно-лагових моделей за цим методом.

Для таких моделей найбільш придатні метод Альмона, метод групового врахування аргументів, та прогнозування нейронимі мережами [3].

Але метод Альмона розроблений для простого регрейсійного аналізу економетричних показників, у якому коефіцієнти β_{t-i} згідно з теоремою Вейерштрасса можна апроксимувати поліномом відповідного ступеня від $-i$ – величини тимчасового лага, має суттеві недоліки: а саме виникає проблема мультиколінеарності змінних та метод не пристосовано для множинного регресійного аналізу. Нейронна мережа придатна для прогнозування ізольованних часових рядів. При прогнозуванні взаїмозалежних рядів число показників обмежено внаслідок ускладнення мережі, та ми не отримуємо регресійної моделі. Крім того, добрі результати були отримані для монотонних моделей.

Тому досліджено метод групового урахування аргументів (МГУА) [1], його чітка та нечітка модифікації [2]. Чіткий, нечіткий алгоритми метода групового урахування аргументів модифіковано для авторегресійних та дистрибутивно лагових моделей.

Застосування чіткої модифікації методу групового урахування аргументів для прогнозування авторегресійних та дистрибутивно-лагових моделей.

Традиційні методи статистичного аналізу даних та прогнозування часових рядів, базуються на математичному апараті, які може бути використаний лише при достатньому обсязі статистичних даних. Часто обсяг даних є недостатнім для пошуку залежностей для великої кількості вхідних факторів. Найкращим виходом з цієї ситуації є використання метода групового урахування аргументів (МГУА), запропонованого академіком О. Івахненко [1].

В даної роботі запропоновано використання метода групового урахування аргументів для множинного регресійного аналізу авторегресійних та дистрибутивно лагових моделей. Було обрано багаторядні поліноміальні моделі МГУА.

Багаторядні алгоритми МГУА застосовуються для рішення некоректних чи недовизначених задач моделювання, тобто у випадку, коли число точок у таблиці дослідних даних не більше числа аргументів, що входять у синтезовану модель. Методи регресійного аналізу в цьому випадку незастосовні, тому що не дають можливості побудови єдиної моделі, адекватної процесу.

Вибірка даних розбиваються на дві: перевірочну та навчальну вибірки. Вважаємо, що початковий склад аргументів, з якого починається процедура багаторядної селекції моделі процесу, будується на так званому нульовому ряді алгоритму, що організується по-різному в поліноміальний і гармонійних алгоритмах.

У випадку авторегресійних моделей до аналогічного виду зводиться модель, отримана за допомогою перетворення усіх вхідних змінних, запізнювань вихідних змінних і заданих нелінійних функцій від них.

$$y_k^{(1)} = a_0 + a_i x_i + a_j x_j + a_{ij} x_i x_j + a_{ii} x_i^2 + a_{jj} x_j^2 + a_{ij+1} * y_{k-1}. \tag{2}$$

Число часткових описів 1-ого ряду дорівнює $M = n(n-1)/2$.

Для дистрибутивно лагової моделі використовують часткові описи квадратичного типу:

$$y_k^{(1)} = a_0 + a_i x_i(t) + a_j x_j(t-1) + a_{ij} x_i(t) x_j(t-1) + a_{ii} x_i^2(t) + a_{jj} x_j^2(t-1) \tag{3}$$

до j-номера тих змінних, для яких автакореляційна функція $\mathrm{corr}(x_j, y, 1) > 0,5$. Багаторядні алгоритми, як правило, працюють за наступною схемою:

1-ий ряд – на основі даних таблиці спостережень будуються часткові описи від усіх попарних комбінацій початкових даних (перепозначених) аргументів, що наближають по МНК вихідну змінну y:

$$y_1 = f_1(x_1, x_2),\ y_2 = f_2(x_1, x_3),\ ...,\ y_k = f_k(y(t-1), x_n). \tag{4}$$

З цих моделей вибирається деяке число кращих за зовнішнім критерієм селекції.

2-ий ряд – отримані змінні приймаються як аргументи – входи другого ряду, і знову будуються всі часткові описи від двох аргументів:

$$z_1 = \varphi_1(y_1, y_2),\ z_2 = \varphi_2(y_1, y_3),\ ...,\ z_l = \varphi_l(y_{F_1-1}, y_F). \tag{5}$$

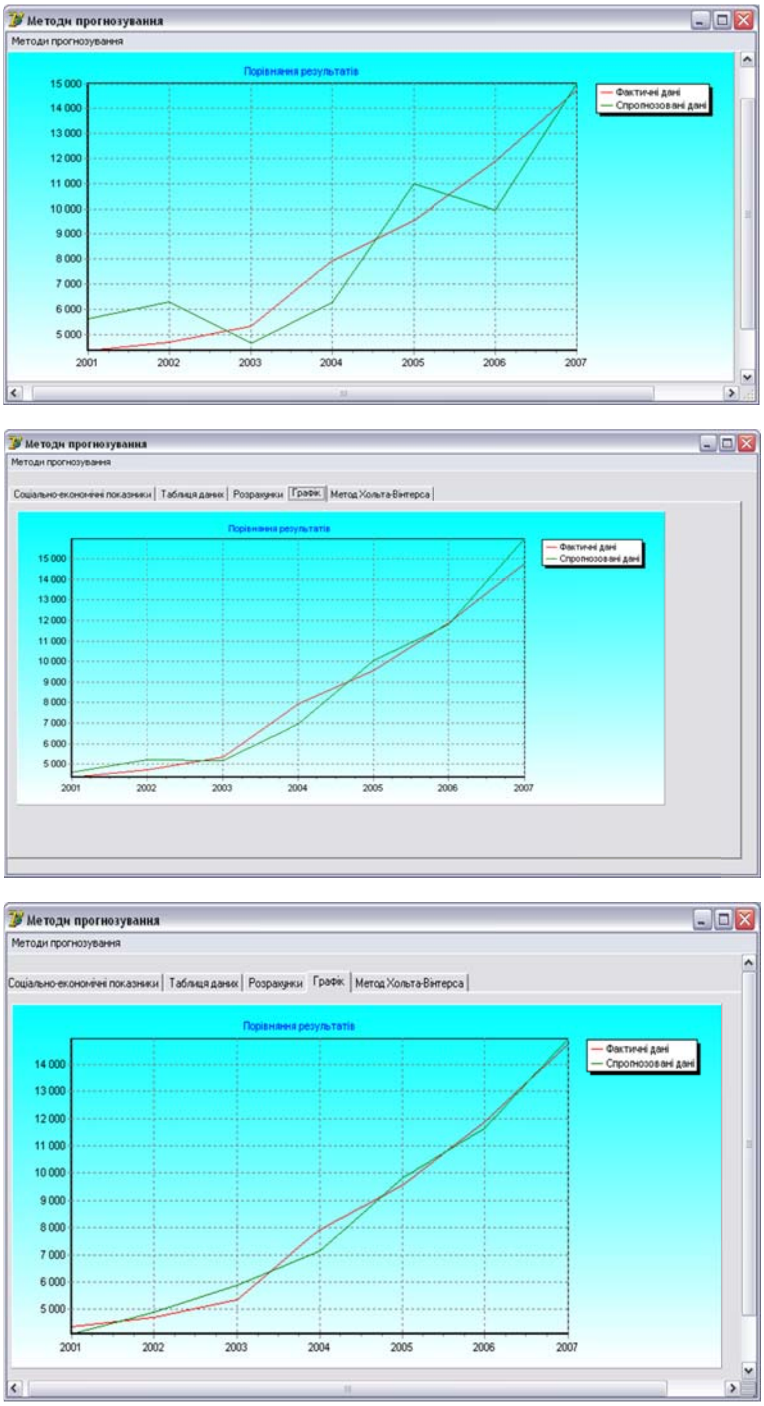
З них за зовнішнім критерієм відбирається F_2 кращих моделей у якості змінних наступного ряду і т. д. Ряди нарощуються доти, поки знижується значення зовнішнього критерію. Виключивши проміжні змінні після останову алгоритму, одержимо модель, у початковому просторі описів число коефіцієнтів у якій значно перевищує число точок

Існує два підходи при виборі часткових моделей точнісний та робастний [2].

Подробний алгоритм чіткої модифікації МГУА для авторегресійних та дистрибутивно лагових моделей наведен в [3]

Було розроблено інформаційну аналітичну систему, яка реалізує багаторядний алгоритм МГУА для авторегресійної моделі. Нижче наведено порівняльний графік прогнозованих та реальних значень (для

ВВП Миколаївської обл.) Залежність не лише від попередніх значень вихідної змінної, а й від 5 показників. Обсяг навчальної та перевірочної вибірок – 7 точок.



а)
б)
в)

Рис 1. Порівняння прогнозованих та точних значень при прогнозуванні за допомогою МГУА:
а) багаторядний алгоритм; б) багаторядний алгоритм та авторегресійна модель;
в) багаторядний алгоритм та дистрибутивно-лагова модель

Застосування нечіткої модифікації методу групового урахування аргументів для прогнозування авторегресійних та дистрибутивно-лагових моделей.

Нечіткий МГУА (НМГУА), розроблений в Інституті прикладного системного аналізу Ю. П. Зайченко [2]. Ця модифікація вдосконалює чіткий метод. Вона відкидає обмеження на вродженість вхідної матриці даних на відміну від МГУА. ***Переваги нечіткого МГУА [2]:***

– відсутня проблема поганої обумовленості матриць, оскільки нема необхідності застосовувати МНК, а задача лінійного програмування завжди має розв'язок;

– отримуємо інтервальну, а не точкову оцінку прогнозованої величини, що дозволяє оцінити точність одержуваних прогнозів;

– застосування нечіткого МГУА в задачах прогнозування економічних процесів зі складною динамікою та невідомим функціональним взаємозв’язком між процесами є цілком аргументованим і дозволяє отримати відносно високу точність пронозу;

– результати прогнозування на НМГУА практично не залежать від типу функції приналежності.

– відсутність похибок, які виникають у МНК і пов’язані з авторегресію випадкових величин, що особливо важливо для авторегресійних моделей.

Розглянемо метод оцінювання лінійної авторегресійної моделі. Нехай є M спостережень $n + 1$ змінної, причому n з них – незалежні величини, а $(n + 1)$-а залежить від інших, і цю залежність $(x_1, \ldots, x_n, y)$ ми намагаємося визначити. При цьому $x_1 = (x_{i1}, \ldots, x_{iM})$ і $y_1 = (y_i, \ldots, y_M)$ – вхідні і вихідні вектори точок спостереження. Тоді оціночна лінійна інтервальна модель для часткової моделі НМГУА має вигляд:

$$Y_k^* = A_{kj}^* z_{kj}^r. \tag{6}$$

У нашому випадку змінні z_i зв’язані зі змінними x_i і x_j для відповідної часткової моделі НМГУА так:

– для дистрибутивно-лаговой моделлі

$$z_0 = 1,\ z_1 = x_{ki},\ z_2 = x_{kj},\ z_3 = x_{k-1,j}. \tag{7}$$

– для авторегресійної моделлі

$$z_0 = 1,\ z_1 = x_{ki},\ z_2 = y_{k-1},$$

де k – номер спостережіння.

Для побудови часткової моделі НМГУА використовувалася лінійна інтервальна регресійна модель, що задається таким чином:

$$Y = B_1 z_1 + B_2 z_2 + \ldots + B_n z_n, \tag{8}$$

де z_i – деякі відомі змінні, B_i – інтервали, які можна задати трикутними нечіткими числами (нечітке число трикутного вигляду) і записати таким чином у вигляді центра інтервалу a_i і його ширини c_i: $B_i = (a_i, c_i)$.

Виходячи з цього, Y є нечітким числом, і його параметри можна розрахувати так [2]:

$$Y = \left(\sum a_i z_i,\ \sum c_i \mid z_i \mid\right),$$

$$a_y = \sum_{i=1}^{n} a_i z_{ik} = a^T z, \qquad c_y = \sum_{i=1}^{n} c_i |z_{ik}| = c^T |z|.$$

Побудова робиться з урахуванням наступних вимог:

– Задані значення y_k, що спостерігаються, включаються в оціночний інтервал Y_k^*.

– Ширина оціночного інтервалу повинна бути мінімальною.

Тобто, для того, щоб інтервальна оцінка була коректною, необхідно, щоб дійсне значення залежної величини навчальної вибірки належало інтервалу, який визначається формулами:

$$\begin{cases} a^T z - c^T \mid z \mid \le Y; \\ a^T z + c^T \mid z \mid \ge Y. \end{cases}$$

Задача лінійного програмування для k-ої точки спостереження:

$$\min(c_0 + c_1 \cdot |x_i| + c_2 \cdot |x_j|),$$

$$a_0 + a_1 \cdot x_i + a_2 \cdot x_j - (c_0 + c_1 \cdot |x_i| + c_2 \cdot |x_j|) \le Y_k,$$

$$a_0 + a_1 \cdot x_i + a_2 \cdot x_j + (c_0 + c_1 \cdot |x_i| + c_2 \cdot |x_j|) \ge Y_k,\ \mathrm{c_p} \ge 0, \mathrm{p} = \overline{0{,}5}.$$

Виходячи з цього, знаючи значення змінних x_i та величини у, отриманих в M спостереженнях, ми переходимо до наступної задачі пошуку коефіцієнтів моделі (для всіх точок спостереження): [2]

$$\min\left(c_0 + c_1 \cdot \sum_{k=1}^{M} |x_i| + c_2 \cdot \sum_{k=1}^{M} |x_j|\right), \tag{9}$$

$$a_0 + a_1 \cdot x_i + a_2 \cdot x_j - (c_0 + c_1 \cdot |x_i| + c_2 \cdot |x_j|) \le Y_k,$$

$$a_0 + a_1 \cdot x_i + a_2 \cdot x_j + (c_0 + c_1 \cdot |x_i| + c_2 \cdot |x_j|) \ge Y_k,$$

$$k = \overline{1, M},\ c_p \ge 0,\ p = \overline{0, 5}.$$

де k – номер спостереження, з якого беруться дані.

Оскільки дана задача є задачею лінійного програмування, для її рішення перейдемо до двоїстої задачі [2]:

$$\max\left(\sum_{k=1}^{M} Y_k \cdot \delta_k - \sum_{k=1}^{M} Y_k \cdot \delta_{k+M}\right), \tag{10}$$

$$\left.\begin{aligned} &\sum_{k=1}^{M} \delta_k - \sum_{k=1}^{M} \delta_{k+M} = 0 \\ &\sum_{k=1}^{M} x_{ki} \cdot \delta_k - \sum_{k=1}^{M} x_{ki} \cdot \delta_{k+M} = 0 \\ &\sum_{k=1}^{M} x_{kj} \cdot \delta_k - \sum_{k=1}^{M} x_{kj} \cdot \delta_{k+M} = 0 \end{aligned}\right\},$$

$$\left.\begin{aligned}&\sum_{k=1}^{M}\delta_k+\sum_{k=1}^{M}\delta_{k+M}\le M\\&\sum_{k=1}^{M}|x_{ki}|\cdot\delta_k+\sum_{k=1}^{M}|x_{ki}|\cdot\delta_{k+M}\le\sum_{k=1}^{M}|x_{ki}|\\&\sum_{k=1}^{M}|x_{kj}|\cdot\delta_k+\sum_{k=1}^{M}|x_{kj}|\cdot\delta_{k+M}\le\sum_{k=1}^{M}|x_{kj}|\end{aligned}\right\},$$

$$\delta_k\ge 0;\ k=\overline{1,\,2M}.$$

Запропоновано вирішити двоїсту задачу двоїстим симплекс-методом. Ми знайдемо оптимальні значення шуканих змінних a_i та c_i ($i=\overline{0,5}$), а разом з тим визначимо шукану модель математичної залежності. Це дозволить скоротити нечіткий МГУА та нечіткий МГУА з нечіткими входами та отримати покращені результати. Ця задача завжди має рішення, оскільки при $\delta^k=0$, $k=\overline{1,2\cdot M}$ всі обмеження виконуються.

Для нечіткого алгоритму МГУА з нечіткими входами розглядається [2] частковий опис вигляду лінійної моделі $f(x_i,x_j)=A_0+A_1x_i+A_2x_j$. Тоді $z_0=1$, $z_1=x_{ki}$, $z_2=y_{k-1}$, A_i – нечіткі числа трикутної форми, які описуються трійкою параметрів $A_i=(\underline{A_i},a_i,\overline{A_i})$, де a_i – центр інтервалу; $\overline{A_i}$ – його верхня межа; $\underline{A_i}$ – нижня межа. $\underline{A_i}=a_i-c_i$; $\overline{A_i}=a_i+c_i$, c_i – ширина інтервалу, $c_i\ge 0$; Z_i – також нечіткі числа трикутної форми, які задаються параметрами $(\underline{Z_i},\breve{Z}_i,\overline{Z_i})$, $\underline{Z_i}$ – нижня межа, $\breve{Z}_i$ – центр, $\overline{Z_i}$ – верхня межа нечіткого числа. Тоді y – нечітке число, параметри якого визначаться таким чином: центр інтервалу $\breve{y}=\sum_i a_i*\breve{Z}_i$ відхилення в лівій частині ФП: $\breve{y}-\underline{y}=\sum_i|a_i|*(\breve{Z}_i-\underline{Z}_i)+c_i|\breve{Z}_i|$ звідки нижня межа інтервалу: $\underline{y}=\sum_i a_i*\breve{Z}_i-$ $-|a_i|*(\breve{Z}_i-\underline{Z}_i)-c_i|\breve{Z}_i|$, відхилення в правій частині ФП: $\overline{y}-\breve{y}=\sum_i|a_i|*(\overline{Z}_i-\breve{Z}_i)+c_i|\breve{Z}_i|$, звідки верхня границя інтервалу: $\overline{y}=\sum_i|a_i|*(\overline{Z}_i-\breve{Z}_i)+c_i|\breve{Z}_i|+a_i*\breve{Z}_i$. Тоді математична модель запишеться в наступному вигляді [2]:

$$\min_{a_i,c_i}\left((2Mc_0+|a_1|\sum_{k=1}^{M}(\overline{x}_{ik}-\underline{x}_{ik})+2c_1\sum_{k=1}^{M}|\breve{x}_{ik}|+|a_2|\sum_{k=1}^{M}(\overline{x}_{jk}-\underline{x}_{jk})+2c_2\sum_{k=1}^{M}|\breve{x}_{jk}|\right.,$$

за умов

$$a_0+a_1\underline{x}_{ik}+a_2\underline{x}_{jk}-c_0-c_1|\breve{x}_{ik}|-c_2|\breve{x}_{jk}|\le y_k,\tag{11}$$

$$a_0+a_1\overline{x}_{ik}+a_2\overline{x}_{jk}+c_0+c_1|\breve{x}_{ik}|+c_2|\breve{x}_{jk}|\ge y_k,\ c_i\ge 0,\ i=\overline{0,5}.$$

Як видно, це завдання є завданням лінійного програмування, але оскільки немає обмежень позитивності для змінних, то для її вирішення переходимо до подвійної задачі, вводячи подвійні змінні $\{\delta_k\}$ і $\{\delta_{k+M}\}$.

Запишемо подвійну задачу [2]:

$$\max\left(\sum_{k=1}^{M}y_k\delta_{k+M}-\sum_{k=1}^{M}y_k\delta_k\right)\tag{12}$$

за умов

$$\left.\begin{aligned}&\sum_{k=1}^{M}\delta_{k+M}-\sum_{k=1}^{M}\delta_k=0\\&\sum_{k=1}^{M}\overline{x}_{ik}\delta_{k+M}-\sum_{k=1}^{M}\underline{x}_{ik}\delta_k=\sum_{k=1}^{M}(\overline{x}_{ik}-\underline{x}_{ik})\\&\sum_{k=1}^{M}\overline{x}_{jk}\delta_{k+M}-\sum_{k=1}^{M}\underline{x}_{jk}\delta_k=\sum_{k=1}^{M}(\overline{x}_{jk}-\underline{x}_{jk})\end{aligned}\right\},$$

$$\left.\begin{aligned}&\sum_{k=1}^{M}\delta_{k+M}+\sum_{k=1}^{M}\delta_k\le 2M\\&\sum_{k=1}^{M}|\breve{x}_{ik}|\delta_{k+M}+\sum_{k=1}^{M}|\breve{x}_{ik}|\delta_k\le 2\sum_{k=1}^{M}|\breve{x}_{ik}|\\&\sum_{k=1}^{M}|\breve{x}_{jk}|\delta_{k+M}+\sum_{k=1}^{M}|\breve{x}_{jk}|\delta_k\le 2\sum_{k=1}^{M}|\breve{x}_{jk}|\end{aligned}\right\},$$

$$\delta_k \geq 0,\ \delta_{k+M} \geq 0,\ k = \overline{1, M}.$$

Вирішивши двоїсту задачу двоїстим симплекс-методом ми знайдемо оптимальні значення шуканих змінних a_i та c_i ($i = \overline{0{,}5}$), а разом з тим визначимо шукану модель математичної залежності.

На рисунках 2, 3 представлені блок-схеми алгоритмів МГУА, що реалізовані у інтелектуально-інформаційної системі.

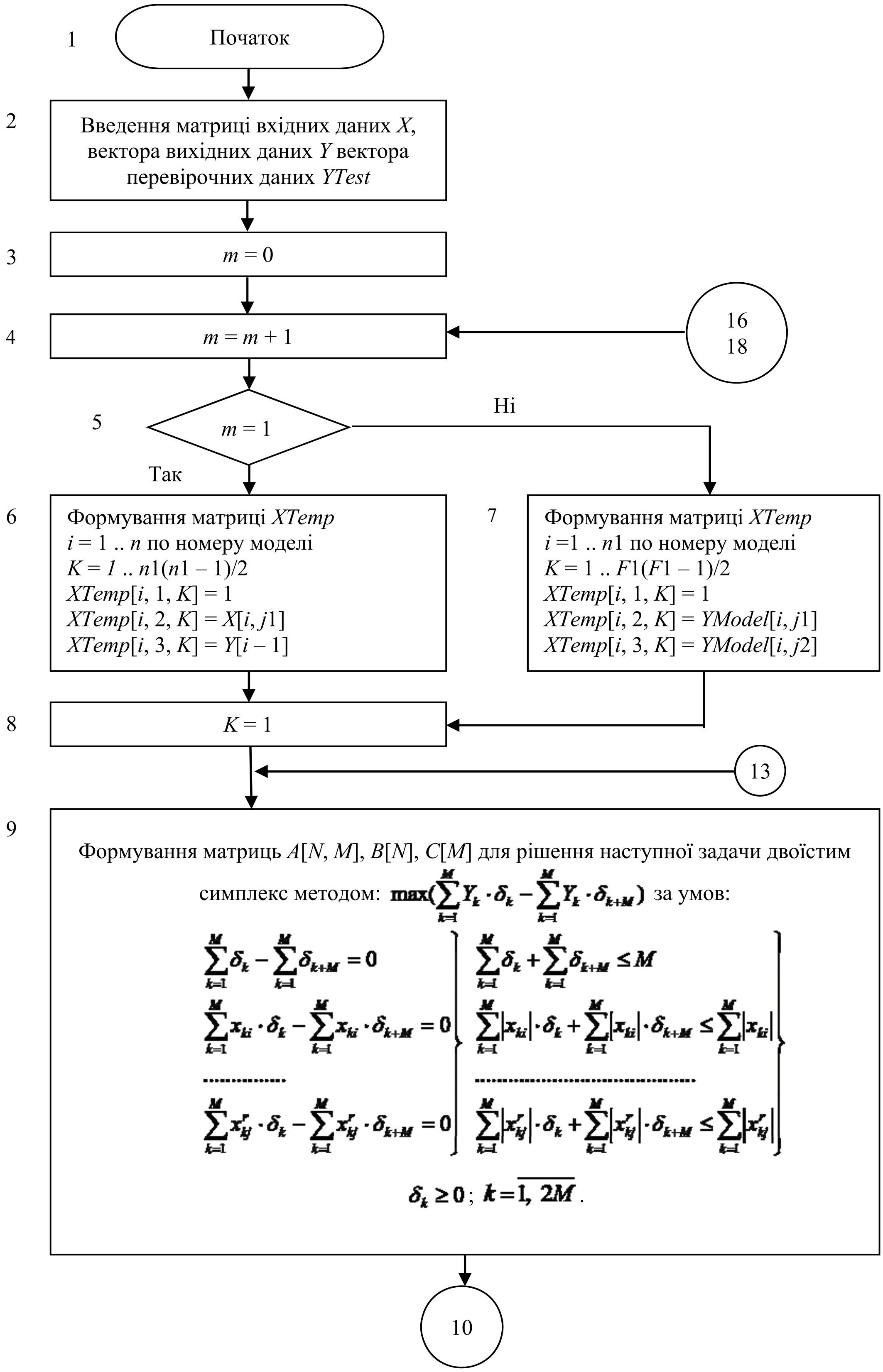


Рис. 2. Блок-схема нечіткого алгоритму МГУА для прогнозування авторегресійних моделей

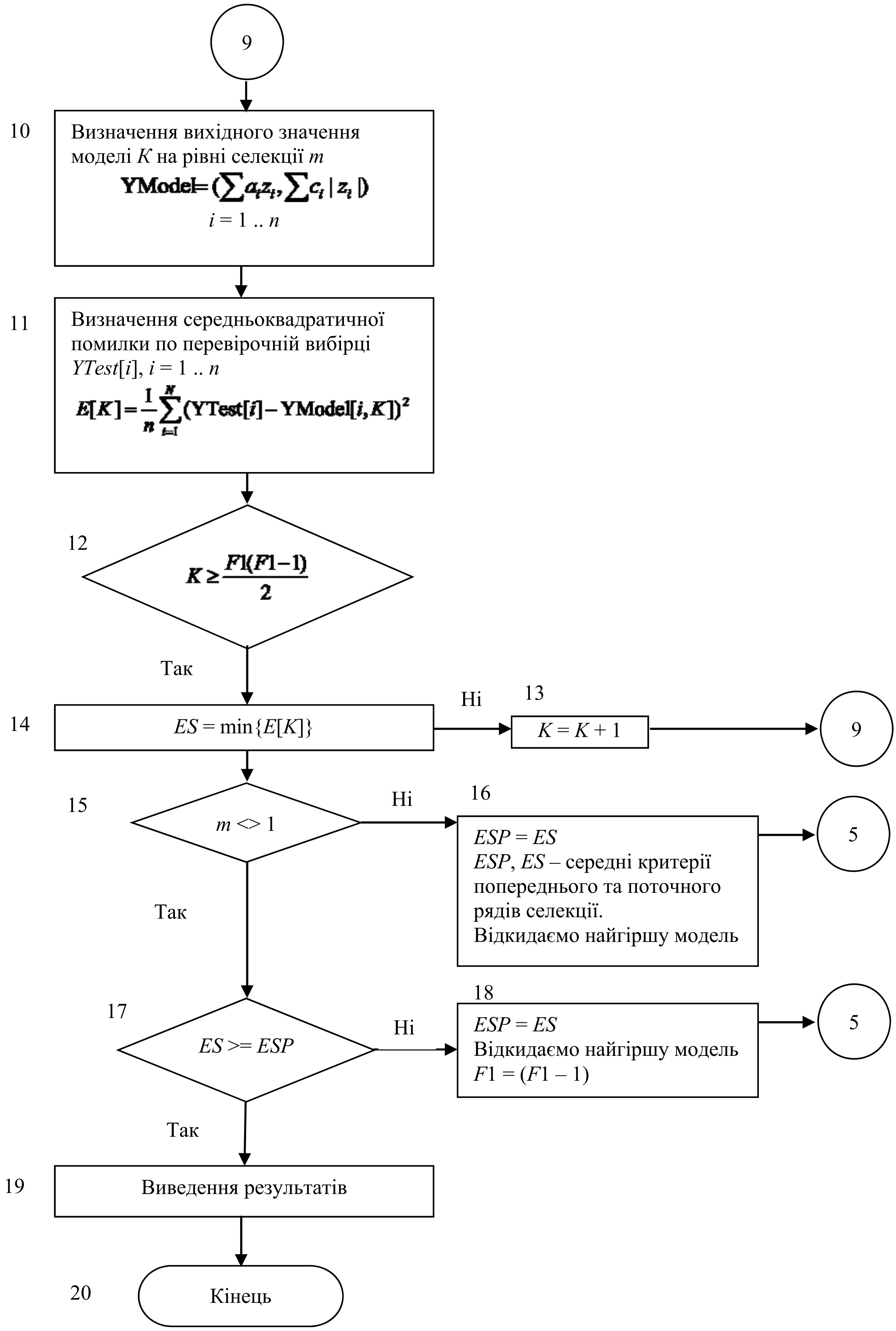


Рис. 2. Аркуш 2

Порівняльний аналіз чітких та нечітких алгоритмів МГУА для авторегресійних моделей

Для проведення порівняльного аналізу чітких та нечітких алгоритмів МГУА для авторегресійних моделей була обрана задача прогнозування макроекономічних показників України та використано значення середньоквадратичної помилки. У методах використовувався критерій регулярності послідовності (точнісний критерій) та частковий опис типу $A_{00} + A_{01} \cdot x_1 + A_{02} \cdot x_2$.

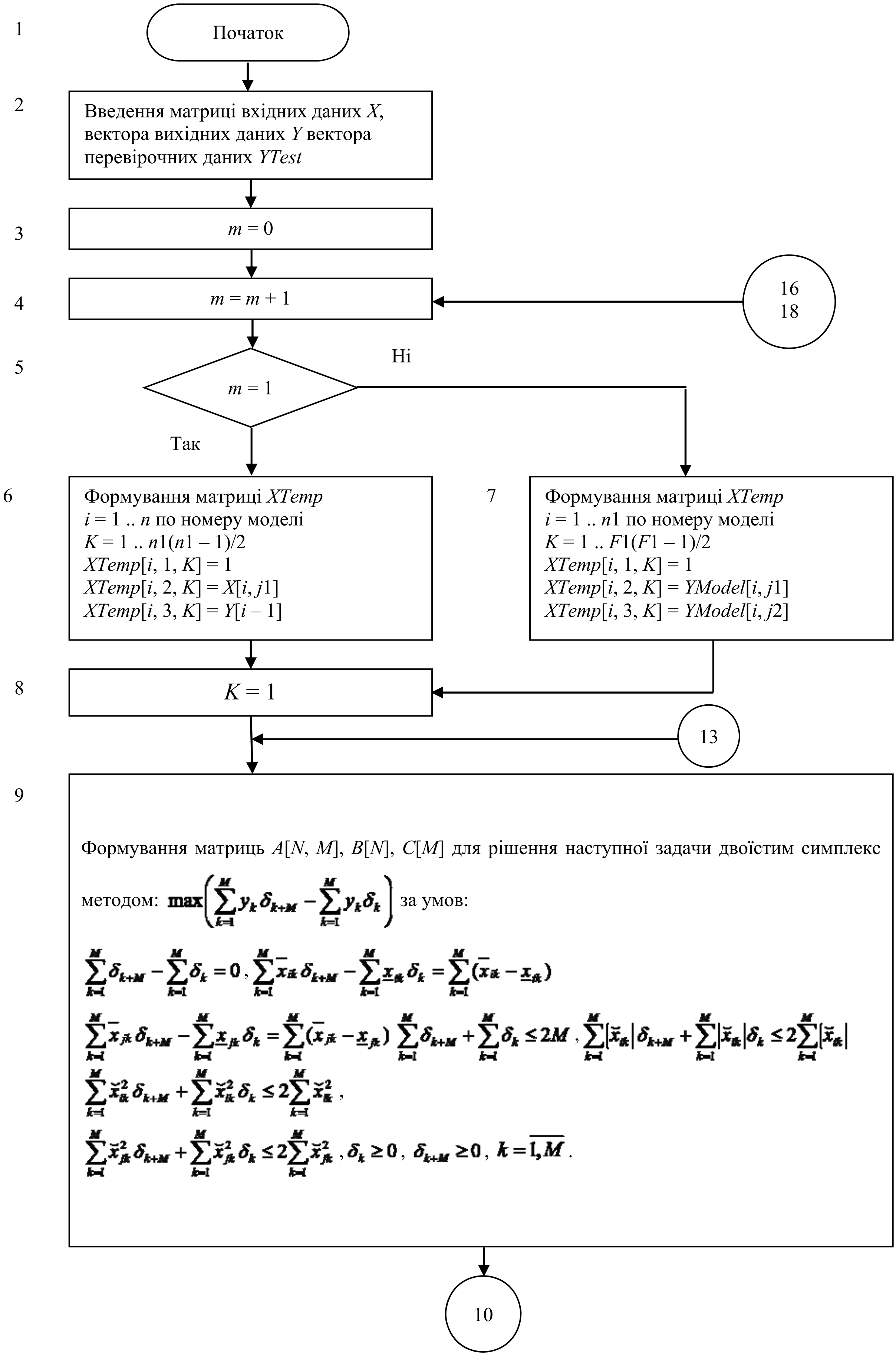


Рис. 3. Блок-схема нечіткого алгоритму МГУА з нечіткими вхідними даними для прогнозування авторегресійних моделей

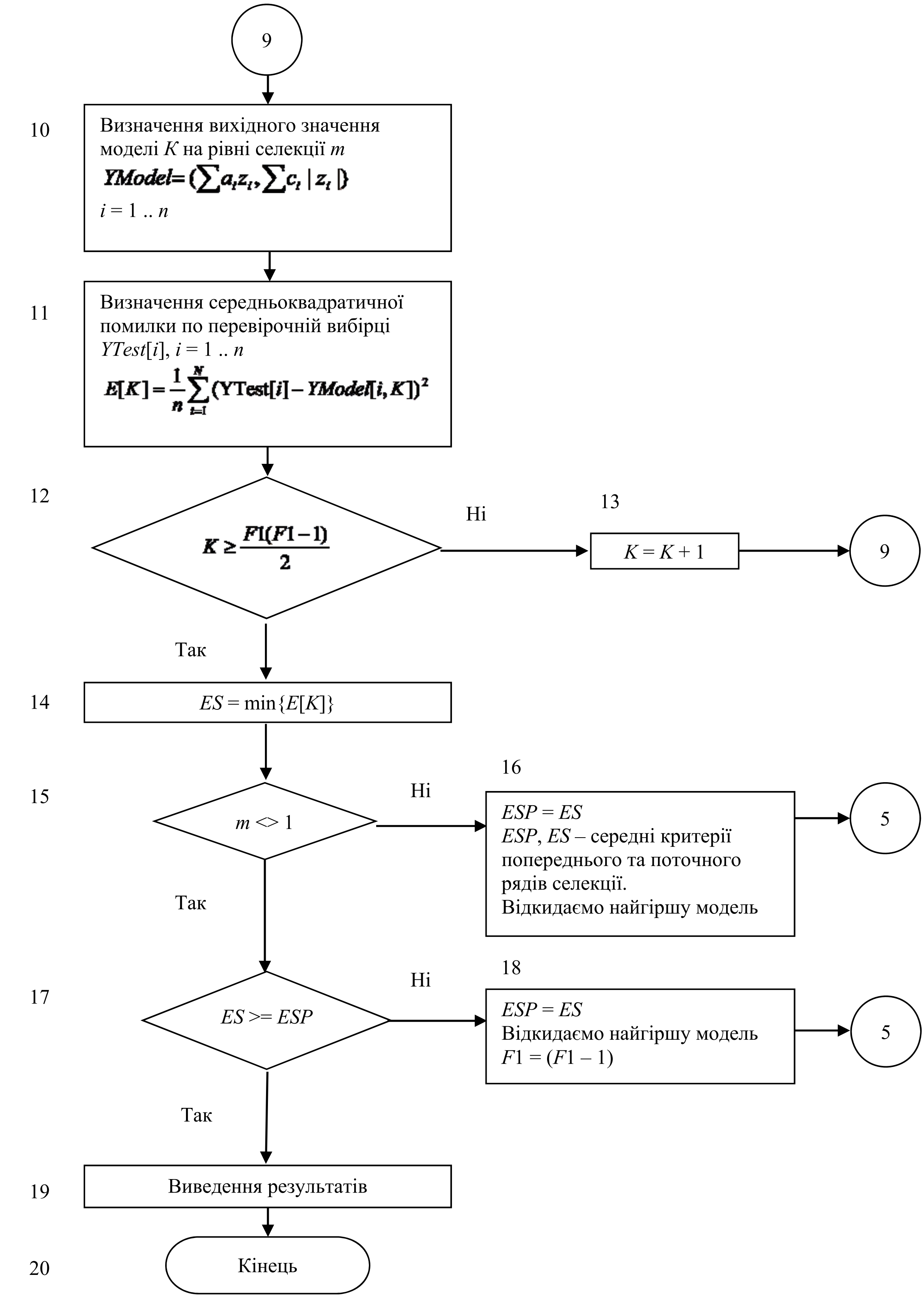


Рис. 3. Аркуш 2

а) Прогноз був виконаний за допомогою багаторядного поліноміального алгоритму МГУА та нечіткого МГУА. В якості вихідної величини, яка прогнозується, було обрано ІСЦ (CPI) – індекс споживчих цін (customer price index). В якості вхідних змінних були обрані наступні змінні: індекси споживчих цін попереднього року та цін виробників, промислове виробництво, валовий внутрішній продукт, грошовий агрегат M_2, доходи населення, середньомісячна заробітна плата, обмінний курс.

Чіткий алгоритм МГУА СКП 0,000824758

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхілення
544,193336	544,020020	0,0288965	0,000835008
535,084983	537,210022	0,0286935	0,000823315
534,363454	532,599976	0,0286317	0,000819777
533,458266	533,270020	0,0286872	0,000822955
530,918608	531,969971	0,0286834	0,000822738

Нечіткий алгоритм МГУА СКП 0,000823284

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхілення
544,455236	544,020020	0,0286209	0,00082162
536,941417	537,210022	0,0287221	0,000824957
533,026056	532,599976	0,0286783	0,000823268
533,003385	533,270020	0,0286359	0,000822482
532,395547	531,969971	0,028707	0,000824092

Рис. 4. Моделювання чітким багаторядним та нечітким МГУА

Отже, за результатами роботи алгоритмів, середньоквадратична помилка:
Багаторядний поліноміальний алгоритм МГУА:
Величина СКП: 0,000824758.
Нечіткий МГУА:
Величина СКП: 0,000823284.
б) Прогноз був виконаний за допомогою чіткого алгоритму МГУА з послідовним виділенням трендів та нечіткого МГУА з нечіткими вхідними змінними Отриманий результат:

Чіткий алгоритм МГУА СКП 0,000823873

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхілення
544,051379	544,020020	0,0287044	0,000823944
537,117572	537,210022	0,0287029	0,000823857
532,583169	532,599976	0,0287134	0,00082446
533,347917	533,270020	0,0286987	0,000823617
530,527706	531,969971	0,0286964	0,000823486

Нечіткий алгоритм МГУА СКП 0,0008231

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхілення
544,509638	544,020020	0,0286163	0,000821357
536,833975	537,210022	0,0287209	0,000824892
533,079316	532,599976	0,0286737	0,000823004
532,896731	533,270020	0,0286348	0,000822418
532,448744	531,969971	0,0287024	0,000823828

Рис. 5. Моделювання МГУА з послідовним виділенням трендів та нечітким МГУА

Отже, за результатами роботи алгоритмів, середньоквадратична помилка:
Чіткий алгоритму МГУА з послідовним виділенням трендів:
Величина СКП: 0,000823873.
Нечіткий МГУА з нечіткими вхідними змінними:
Величина СКП: 0,0008231.

Як видно з наведених прикладів при використанні нечітких алгоритмів результат є більш точним. Це відбувається тому, що в нечіткому МГУА використовуємо адаптацію коефіцієнтів знайденої нечіткої моделі за поточними даними, що дозволяє підвищити точність прогнозування на 15-20 % та відсутня проблема поганої обумовленості матриць, оскільки нема необхідності застосовувати МНК, а задача лінійного програмування завжди має розв'язок, що також забезпечує більш стабільний результат та високу точність прогнозу.

У залежності від параметрів, що варіюються в експериментах, можна зробити наступні висновки:

– велика кількість вхідних параметрів безсумнівно підвищує якість прогнозів як в чіткому, так і в нечіткому алгоритмах МГУА;

– при однаковій кількості вхідних параметрів зміна співвідношення між навчальною та перевірочною вибірками (збільшення навчальної та зменшення перевірочної) веде до погіршення результатів прогнозу. Зокрема, нечіткий метод дає ширший інтервал значень, чіткий – більш згладжену криву;

– застосування нечіткого МГУА в задачах прогнозування економічних процесів зі складною динамікою та невідомим функціональним взаємозв'язком між процесами дозволяє отримати відносно високу точність прогнозу.

– порівнювати МГУА з іншими методами прогнозування неможливо, оскільки останні не можуть вирішити задачу при обмеженій кількості статистичних даних, а також вино лише прогнозують у часі, не показуючи та не враховуючи залежності між вхідними показниками.

Порівняльний аналіз нечіткого та нечіткого з нечіткими входами алгоритмів МГУА

Для проведення порівняльного аналізу нечіткого та нечіткого з нечіткими входами алгоритмів МГУА використаємо значення середньоквадратичної помилки, що обраховується при отриманні результатів роботи кожного з цих алгоритмів. У методах використовувався критерій регулярності послідовності (точнісний критерій) та частковий опис типу $A_{00} + A_{01} \cdot x_1 + A_{02} \cdot x_2$. Було проведено два дослідження: на авторегресійної та на дистрибутивно-лагової моделі.

а) Тип моделі – авторегресійна Нечіткий алгоритм:

Нечіткий алгоритм МГУА СКП 0,000823284

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхилення
544,455236	544,020020	0,0286209	0,00082162
536,941417	537,210022	0,0287221	0,000824957
533,026056	532,599976	0,0286783	0,000823268
533,003385	533,270020	0,0286359	0,000822482
532,395547	531,969971	0,028707	0,000824092

б) Нечіткий з нечіткими вхідними даними:

Нечіткий алгоритм МГУА СКП 0,0008231

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхилення
544,509638	544,020020	0,0286163	0,000821357
536,833975	537,210022	0,0287209	0,000824892
533,079316	532,599976	0,0286737	0,000823004
532,896731	533,270020	0,0286348	0,000822418
532,448744	531,969971	0,0287024	0,000823828

Рис. 6. Моделювання авторегресійної моделі за допомогою нечіткого МГУА з нечіткими вхідними даними

Отже, за результатами роботи алгоритмів, середньоквадратична помилка:
Нечіткий МГУА:
Величина СКП: 0,000823284.
Нечіткий МГУА з з нечіткими вхідними даними:
Величина СКП: 0,0008231.

в) Тип моделі – дистрибутивно-лагова. Нечіткий алгоритм.

Нечіткий алгоритм МГУА СКП 0,00082262

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхилення
543,666407	544,020020	0,0286204	0,000821594
537,451767	537,210022	0,0286938	0,000823335
532,253786	532,599976	0,0286778	0,000823242
533,509992	533,270020	0,0286077	0,000820865
531,624191	531,969971	0,0287066	0,000824066

Рис. 7. Моделювання дистрибутивно-лагової моделі за допомогою нечіткого МГУА з нечіткими вхідними даними

Нечіткий алгоритм МГУА СКП 0,000822613

Прогнозоване зн.	Точне зн.	Відхилення	Квадрат відхилення
543,693608	544,020020	0,0286199	0,000821561
537,424906	537,210022	0,0286944	0,000823368
532,280416	532,599976	0,0286773	0,000823209
533,483328	533,270020	0,0286083	0,000820898
531,650789	531,969971	0,028706	0,000824033

Рис. 8. Моделювання дистрибутивно-лагової моделі за допомогою нечіткого МГУА з нечіткими вхідними даними

Отже, за результатами роботи алгоритмів, середньоквадратична помилка:
Нечіткий МГУА:
Величина СКП: 0,00082262.
Нечіткий МГУА з з нечіткими вхідними даними:
Величина СКП: 0,000822613.

Як видно з порівняльних таблиць, МГУАНВ на даній задачі та при даних вхідних параметрах продемонстрував кращі значення показників прогнозу на перевірочній вибірці.

Основним досягненням МГУАНВ у порівнянні з НМГУА є зменшення інтервалу ширини прогнозу. Причиною такого досягнення можна вважати те, що інтервал ширини прогнозу у методах НМГУА та МГУАНВ вираховується за двома різними принципами. У НМГУА інтервал з'являється за рахунок нечітких коефіцієнтів моделі, а у МГУАНВ за рахунок нечітких значень коефіцієнтів моделі та вхідних параметрів.

Висновки. Дослідження показали, що метод групового урахування аргументів, якій дозволяє отримати коефіцієнті моделі при недостатньому обсязі даних та має більшу точність прогнозу, є найбільш придатним для авторегресійних моделей та дистрибутивно-лагових моделей. Запропоновано модифікації чіткого МГУА, нечіткого МГУА, та МГУА з нечіткими входами саме для прогнозування авторегресійних та дистрибутивно лагових моделей. Також запропоновано модифікацію нечітких алгоритмів МГУА, що полягає у використанні двоїстого симплекс-методу замість звичайного симплекс-методу, що дозволило скоротити нечіткі МГУА та отримати покращені результати. Усі модифікації МГУА для авторегресійних та дистрибутивно лагових моделей демонструють високу точність прогнозу. Найкращу точність мають нечіткий МГУА, та МГУА з нечіїткими входами. Крім того, їх перевага над чітким МГУА у тому, що вони не використовують МНК та не чутливі до поганої обумовленості матриці вхідних факторів і автокореляції випадкових величин, що так важливо для для авторегресійних моделей.

ЛІТЕРАТУРА

1. Ивахненко А. Г. Принятие решений на основе самоорганизации / А. Г. Ивахненко, Ю. П. Зайченко, В. Д. Димитров. – М. : «Сов. радио», 1976. – 280 с.
2. Зайченко Ю. П. Нечеткие модели и методи в интеллектуальних системах. – К. : «Слово», 2008. – 333 с.
3. Кравець І. О. Дослідження методів статистичного та інтелектуального аналізу для авторегресійних моделей / І. О. Кравець, Г. О. Афанасьєва // Наукові праці : науково-методичний журнал. – Миколаїв : Вид-во ЧДУ ім. Петра Могили, 2009. – Вип 93.

Стаття надійшла до редколегії 28.03.2011 р.