

УДК 519.688

А.О. Долгіх, О.І. Білобородько, О.Г. Байбуз

Дніпропетровський національний університет імені Олеся Гончара

ЗНАХОДЖЕННЯ ОПТИМАЛЬНИХ ЗНАЧЕНЬ ПАРАМЕТРІВ АДАПТИВНИХ МОДЕЛЕЙ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ З ВИКОРИСТАННЯМ ГЕНЕТИЧНОГО АЛГОРИТМУ

Описано алгоритм пошуку оптимальних значень коефіцієнтів згладжування адаптивних моделей прогнозування часових рядів, реалізований авторами. Запропонований алгоритм був застосований до вирішення задачі прогнозування економічних показників. Виконано аналіз результатів ефективності розробленого алгоритму.

Ключові слова: *прогнозування часових рядів, адаптивні моделі прогнозування, генетичний алгоритм з дійсним кодуванням, економічні часові ряди.*

Описан алгоритм поиска оптимальных значений коэффициентов сглаживания адаптивных моделей прогнозирования временных рядов, реализованный авторами. Предложенный алгоритм был применен к решению задач прогнозирования экономических показателей. Выполнен анализ результатов эффективности разработанного алгоритма.

Ключевые слова: *прогнозирование временных рядов, адаптивные модели прогнозирования, генетический алгоритм с вещественным кодированием, экономические временные ряды.*

Optimal values of adaptive forecasting models smoothing factors search algorithm developed by authors has been described. The proposed algorithm has been applied to economical time series forecasting. Analysis of the developed algorithm efficiency has been done.

Keywords: *time series forecasting, adaptive forecasting models, real coded genetic algorithm, economical time series.*

Вступ. Особливої актуальності в різних галузях людської діяльності набуває задача прогнозування. В економіці – для прогнозування щоденних коливань цін на акції, курсів валют, щотижневих та щомісячних обсягів продаж, річних обсягів виробництва тощо. У природничих науках – для прогнозування кількості опадів, природних явищ, стану забруднення водних ресурсів, оцінки деяких біологічних

та біохімічних показників.

Для вирішення задачі прогнозування вдалим рішенням є застосування адаптивних моделей. Ці моделі дозволяють обчислювати прогнозні значення без перерахунку моделі у разі отримання нових даних, не накладають обмежень на розмір ряду та швидко підлаштовуються під динаміку змін рівнів ряду, ґрунтуються на принципі експоненціального згладжування, тим самим враховують «старіння» інформації [1].

Однак результати застосування цих методів сильно залежать від коефіцієнтів згладжування $\beta_1, \beta_2, \beta_3$. При зміні значень цих коефіцієнтів значення середньоквадратичної похибки може істотно зростати. Задача пошуку параметрів $\beta_1, \beta_2, \beta_3$ не має тривіального вирішення. Тому було вирішено реалізувати програмне забезпечення, яке б знаходило оптимальні параметри адаптивних моделей за допомогою генетичного алгоритму, та дослідити, яким чином параметри генетичного алгоритму впливають на якість побудованого прогнозу.

Аналіз літературних даних. Найбільш відомими моделями прогнозування часових рядів є регресійні моделі. У таких моделях ряд подають у вигляді суми або добутку трьох компонент:

$$f(t) = T(t) + S(t) + A(t) + \xi_t,$$

$$f(t) = T(t) * S(t) * A(t) * \xi_t \quad (1.1),$$

де $T(t)$ – лінія тренда, що показує глобальні зміни досліджуваного явища; $S(t)$ – сезонність, яка відображає коливання відносно тренда, обумовлені зовнішніми впливами; $A(t)$ – циклічність (автоколивання) – більш-менш регулярні коливання відносно тренда, обумовлені внутрішньою природою досліджуваного явища.

Задача прогнозування часових рядів у випадку регресійних моделей зводиться до задачі виділення цих трьох компонент. Трендову компоненту представляють у формі лінійної чи нелінійної регресії, параметри якої шукають методом найменших квадратів. Для подання сезонної та циклічної складової використовують один з двох підходів: індекси сезонності або гармонічний аналіз [2]. Для побудови регресійних моделей необхідно мати велику базу даних спостережень. Корисно також мати незалежний набір прикладів, на яких можна перевірити якість роботи моделі. Окрім цього, окремою проблемою при використанні цих моделей є обрання вигляду лінії регресії.

Сьогодні великої популярності набувають методи, які

представляють собою подальший розвиток методу регресійного аналізу, зокрема метод МГУА – метод групового урахування аргументів. Метод реалізує задачу синтезу оптимальних моделей високої складності, адекватної складності досліджуваного об'єкта (тут під моделями розуміється система регресійних рівнянь). Алгоритми МГУА, побудовані за схемою масової селекції, здійснюють перебирання можливих функціональних описів об'єкта. До явних переваг МГУА належать автоматичне формування структури мережі, простота і швидкодія настроювання параметрів, а також можливість «згорнути» налаштовану мережу безпосередньо в явний математичний вираз. Головною проблемою при використанні цього методу є те, що для якісного результату цей метод, як і всі регресійні моделі, вимагає досить великої репрезентативної вибірки та складних математичних обчислень [3].

Другим великим класом є моделі прогнозування, які базуються на принципі згладжування, тобто враховують «старіння» інформації. Згладжування часових рядів проводять з метою виділення присутніх у ряді тенденцій та видалення аномальних значень. Одним з найпростіших і найпоширеніших підходів є експоненційне згладжування, в основі якого лежить розрахунок експоненційних середніх. На основі цього підходу було розроблено багато моделей прогнозування, зокрема адаптивні моделі, головними перевагами яких у порівнянні з регресійними методами є те, що вони не накладають обмежень на розміри ряду та не потребують виконання складних математичних операцій. Ці моделі є основним об'єктом поточного дослідження.

Варто зауважити, що сьогодні існує широкий клас методів, які використовують сучасні методи машинного навчання: нейронні мережі, дерева рішень, марківські моделі для вирішення задачі прогнозування часових рядів.

Основний матеріал. В адаптивних методах часовий ряд подають у вигляді функції

$$u_t = f(a_{1t}, a_{2t}, \dots, a_{pt}, t) + e_t, \quad (1.2)$$

під час побудови якої відслідковують величину відхилень прогнозних значень від значень вихідного ряду. Адаптивні методи прогнозування поділяються на лінійні та сезонні. У моделях лінійного зростання прогноз обчислюють за формулою:

$$u_{t+\tau} = \hat{a}_{1,t} + \hat{a}_{2,t}\tau, \quad (1.3)$$

де τ – кількість кроків прогнозу; $\hat{a}_{1,t}$, $\hat{a}_{2,t}$ – коефіцієнти адаптивної

моделі в момент часу t , $\hat{a}_{1,t}$ – оцінка того, чого досягли на поточному кроці; $\hat{a}_{2,t}$ – приріст на поточному кроці.

До адаптивних моделей лінійного зростання відносять модель Хольта, модель Тейла – Вейджа, модель Брауна та модель Бокса – Дженкінса. Серед сезонних моделей виділяють лінійну адаптивну, лінійну мультиплікативну, експоненційну адитивну та експоненційну мультиплікативну. Вказані моделі відрізняються засобами знаходження параметрів $\hat{a}_{1,t}$, $\hat{a}_{2,t}$. Формули оцінок параметрів за адаптивними лінійними моделями наведено у табл. 1. $\beta_1, \beta_2, \beta_3$ – коефіцієнти згладжування, які приймають значення від 0 до 1.

Таблиця 1 – Адаптивні моделі лінійного зростання

Модель	Формула підрахунку оцінок $\hat{a}_{1,t}$ та $\hat{a}_{2,t}$
Модель Хольта	$\hat{a}_{1,t} = \beta_1 u_t + (1 - \beta_1)(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}),$ $\hat{a}_{2,t} = \beta_2 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \beta_2) \hat{a}_{2,t-1},$
Модель Тейла – Вейджа	$\begin{cases} \hat{a}_{1,t} = \beta_1 u_t + (1 - \beta_1) \hat{u}_t = \hat{u}_t + \beta_1 e_t, \\ \hat{a}_{2,t} = \hat{a}_{2,t-1} + \beta_1 \beta_2 e_t, \\ e_t = u_t - \hat{u}_t \end{cases}$
Модель Брауна	$\begin{cases} \hat{a}_{1,t} = \hat{u}_t + (1 - \beta^2) e_t, \\ \hat{a}_{2,t} = \hat{a}_{2,t-1} + (1 - \beta^2) e_t. \\ e_t = u_t - \hat{u}_t \end{cases}$
Модель Бокса – Дженкінса	$\begin{cases} \hat{a}_{1,t} = \hat{u}_t + \beta_1 e_t + \beta_3 (e_t - e_{t-1}), \\ \hat{a}_{2,t} = \hat{a}_{2,t-1} + \beta_1 \beta_2 e_t. \\ e_t = u_t - \hat{u}_t \end{cases}$

У табл. 2 подано схему обчислень для адаптивних сезонних моделей. У наведених моделях сезонність – це масиви довжиною l g_1, \dots, g_l , f_1, \dots, f_l , де l – період сезонності. Елементи цих масивів обчислюють на кожному кроці, як і коефіцієнти моделей.

Таблиця 2 – Адаптивні моделі сезонного зростання

Модель	Формула підрахунку оцінок $\hat{a}_{1,t}$ та $\hat{a}_{2,t}$
Лінійна адитивна	$\hat{u}_{t+\tau} = \hat{a}_{1,t} + \hat{a}_{2,t} \cdot \tau + \hat{g}_{t+\tau-l}$ $\hat{a}_{1,t} = \beta_1 (u_t - \hat{g}_{t-l}) + (1 - \beta_1) (\hat{a}_{1,t-1} - \hat{a}_{2,t-1})$ $\hat{a}_{2,t} = \beta_2 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \beta_2) \hat{a}_{2,t-1},$ $\hat{g}_t = \beta_3 (u_t - \hat{a}_{1,t}) + (1 - \beta_3) \hat{g}_{t-l}.$
Лінійна мультиплікативна (модель Уінтерса)	$\hat{u}_{t+\tau} = (\hat{a}_{1,t} + \hat{a}_{2,t} \cdot \tau) \cdot \hat{f}_{t+\tau-l}$ $\hat{a}_{1,t} = \beta_1 \frac{u_t}{\hat{f}_{t-l}} + (1 - \beta_1) (\hat{a}_{1,t-1} + \hat{a}_{2,t-1}),$ $\hat{a}_{2,t} = \beta_2 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \beta_2) \hat{a}_{2,t-1},$ $\hat{f}_t = \beta_3 \frac{u_t}{\hat{a}_{1,t}} + (1 - \beta_3) \hat{f}_{t-l}.$
Експоненційна адитивна	$\hat{u}_{t+\tau} = \hat{a}_{1,t} \cdot \hat{r}_{1,t}^\tau + \hat{g}_{t+\tau-l}$ $\hat{a}_{1,t} = \beta_1 (u_t - \hat{g}_{t-l}) + (1 - \beta_1) (\hat{a}_{1,t-1} \cdot \hat{r}_{t-1}),$ $\hat{r}_t = \beta_2 \frac{\hat{a}_{1,t}}{\hat{a}_{1,t-1}} + (1 - \beta_2) \hat{r}_{t-1},$ $\hat{g}_t = \beta_3 (u_t - \hat{a}_{1,t}) + (1 - \beta_3) \hat{g}_{t-l}.$
Експоненційна мультиплікативна	$\hat{u}_{t+\tau} = (\hat{a}_{1,t} \cdot \hat{r}_{1,t}^\tau) \cdot \hat{f}_{t+\tau-l}$ $\hat{a}_{1,t} = \beta_1 \frac{u_t}{\hat{f}_{t-l}} + (1 - \beta_1) (\hat{a}_{1,t-1} \cdot \hat{r}_{t-1}),$ $\hat{r}_t = \beta_2 \frac{\hat{a}_{1,t}}{\hat{a}_{1,t-1}} + (1 - \beta_2) \hat{r}_{t-1},$ $\hat{f}_t = \beta_3 \frac{u_t}{\hat{a}_{1,t}} + (1 - \beta_3) \hat{f}_{t-l}.$

Характерною ознакою усіх адаптивних моделей є те, що на кожному кроці обчислюються нові значення a_{1t}, a_{2t} , які залежать від коефіцієнтів згладжування $0 < \beta_1, \beta_2, \beta_3 < 1$. Для того щоб отримати якісні результати прогнозування, потрібно підібрати такі параметри $\beta_1, \beta_2, \beta_3$, які найбільш підходять для заданого часового ряду. Для вирішення цієї задачі було реалізовано програмне забезпечення, яке б дозволяло знаходити оптимальні значення $\beta_1, \beta_2, \beta_3$ за допомогою генетичного алгоритму.

Потенційне рішення оптимізаційної задачі (параметри $\beta_1, \beta_2, \beta_3$) представлено у вигляді хромосом. У класичних генетичних алгоритмах хромосоми подаються у вигляді бінарного вектора. Однак бінарне подання хромосом тягне за собою певні труднощі при виконанні пошуку в безперервних просторах, які пов'язані з великою розмірністю простору пошуку [4]. Тому для вирішення поставленої задачі було використано генетичний алгоритм з кодуванням на базі дійсних чисел (RGA).

У ході виконання дослідження було реалізовано такі оператори схрещування неперервних генетичних алгоритмів: арифметичний, геометричний, плоский, лінійний, дискретний, евристичний, *BLX* – α та *SBX* кросовери та два види мутації: проста мутація та мутація Міхалевича. Також передбачено можливість обрати одну із стратегій відбору батьків для схрещування: панмісія, аутбридинг, інбридинг, турнірна селекція та метод рулетки. Розроблений програмний продукт дозволяє комбінувати різні оператори відбору, схрещування та мутації та відслідковувати, як при цьому змінюється середньоквадратична похибка прогнозу *MSE*. Користувач може змінювати розмір популяції, ймовірність мутації, кількість епох алгоритму, а також параметри, специфічні для деяких операторів схрещування дійсних векторів.

Прогнозування часових рядів з використанням адаптивних методів та генетичного алгоритму.

Крок 1. На першому кроці генерують початкову популяцію хромосом. Обсяг популяції N задається користувачем. Кожна хромосома представляє собою набір значень $\beta_1, \beta_2, \beta_3$. Для деяких лінійних моделей прогнозування, наприклад, моделі Тейла – Вейджа і Хольта, значення останнього параметра β_3 не використовується. Значення тих параметрів, які не використовується тією чи іншою моделлю, тотожно дорівнюють 0.

Крок 2. Розраховують значення середньоквадратичної похибки прогнозу за формулою:

$$MSE = \frac{\sum_{j=1}^n (\text{ForecastedValue}_j - \text{ActualValue}_j)^2}{n} \quad (1.4)$$

для кожної хромосоми в популяції. Для цього, використовуючи параметри $\beta_1, \beta_2, \beta_3$, за обраною користувачем адаптивною моделлю будують прогнозний ряд. Після цього розраховують значення середньоквадратичної похибки прогнозу MSE . Чим менше значення MSE , тим краще побудована модель підлаштовується під значення заданого ряду.

Крок 3. Обирають дві хромосоми з популяції за тією стратегією, яка задана користувачем, і проводять схрещування хромосом. У результаті цієї операції отримують дві нові особини, які успадковують ознаки батьків. Обоє «батьків» та обоє «нащадків» додають до їх «проміжної» популяції. Відбір батьків та схрещування проводять M разів. M задається користувачем.

Крок 4. Проводять мутацію. Для кожної хромосоми з популяції генерують деяке випадкове число. Якщо воно менше рівня мутації, то проводять мутацію хромосоми за однією з можливих стратегій: проста мутація або мутація Міхалевича. Вид мутації також обирається користувачем.

Крок 5. На цьому кроці проводять селекцію або відбір хромосом до наступного покоління. Для отриманої «проміжної» популяції розраховують значення MSE . Деяка кількість хромосом з найменшим значенням середньоквадратичної похибки прогнозу Q потрапляє до наступної популяції. Задля того щоб попередити завчасне виродження популяції, випадковим чином генерують ще $N - Q$ хромосом та додають їх до нової популяції.

Кроки 2–5 проводять ітеративно, поки не виконається одна з умов зупинки генетичного алгоритму: досягнеться введена користувачем кількість поколінь або алгоритм зійдеться до єдиного рішення.

Після цього отриманий набір значень коефіцієнтів $\beta_1, \beta_2, \beta_3$ використовують для прогнозування наступних значень ряду.

Результати прогнозування з використанням адаптивних методів, налагоджених за допомогою генетичних алгоритмів.

Дані, які підлягають аналізу та прогнозуванню, представляють собою *ask price* на акції компанії AAON inc. в період з 4 січня 2012 року по 7 листопада 2014 року. *Ask price* – це ціна, за якою власник акції готовий її продати. Довжина ряду, $N = 282$. На рис. 1 зображено вхідні дані на графіку.

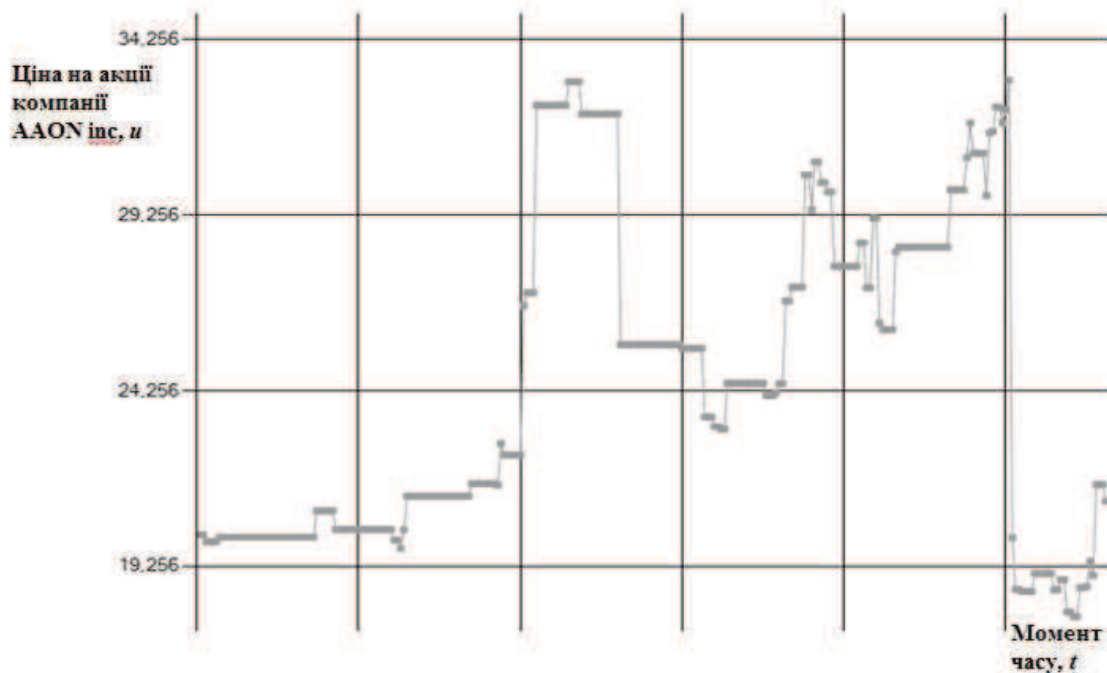


Рисунок 1 – Вхідні дані

У результаті перевірки вхідних даних на випадковість за допомогою критерію Спірмена та серіальної кореляції було встановлено, що введена послідовність не є випадковою, має тенденцію до зростання, а також у ній наявна сезонна складова.

Результати прогнозування досліджуваного ряду за допомогою адаптивних моделей та реалізованого алгоритму пошуку оптимальних значень коефіцієнтів згладжування представлено у табл. 3.

Для налагоджування моделей застосовувався генетичний алгоритм з дійсним кодуванням з такими параметрами: $N = 50$ (розмір популяції), відсоток кращих хромосом, які потрібно залишити популяції, $q = 80\%$, кількість епох (повторень) генетичного алгоритму = 50, рівень мутації = 0,1. Використовувалося значення $K = 3$. Для сезонних моделей значення довжини сезонності $l = 12$. Реальне значення ряду $U_{n+1} = 21,11$.

Таблиця 3 – Результати роботи реалізованого алгоритму пошуку оптимальних значень коефіцієнтів згладжування

Модель	β_1	β_2	β_3	MSE	Побудоване прогнозне значення
Модель Брауна	0,5036	N/A	N/A	1,4139	21,6228
Модель Хольта	0,9841	0,0039	N/A	1,2103	21,0745
Модель Тейла – Вейджа	0,9987	0,0126	N/A	1,2143	21,0616
Модель Бокса – Дженкінса	0,9666	0,0031	0,0479	1,2066	21,048
Лінійна адитивна	0,5836	0,5358	0,3059	1,221	21,1118
Лінійна мультиплікативна	0,0176	0,0065	0,9794	1,2238	21,125
Експоненційна адитивна	0,0033	0,0109	0,9688	1,2216	21,1104
Експоненційна мультиплікативна	0,0061	0,0151	0,9803	1,2206	21,1106

Задача знаходження оптимальних значень параметрів $\beta_1, \beta_2, \beta_3$ може бути розглянута як задача оптимізації в багатовимірному просторі. В даному випадку потрібно знайти такі значення $\beta_1, \beta_2, \beta_3$, при яких MSE приймає мінімальне значення. Тому програма надає можливість побудувати графік поверхні, для якої шукають екстремум (рис. 2). У якості координат X та Y використовують значення β_1, β_2 або β_3 . У якості координати Z виступає значення $1/MSE$. Ділянка, де $1/MSE$ досягає максимального значення, позначена на графіку червоним кольором.

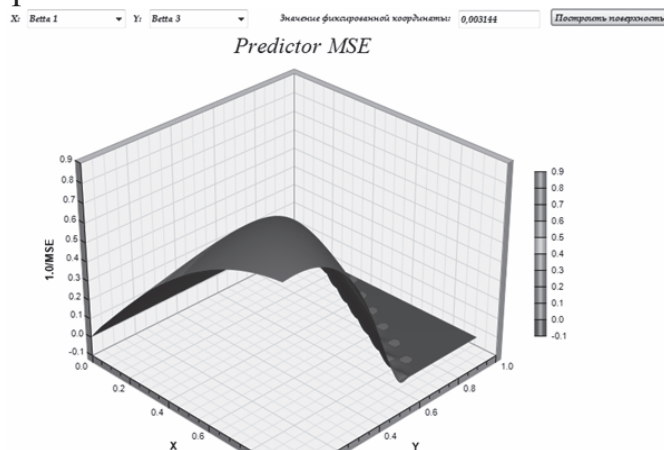


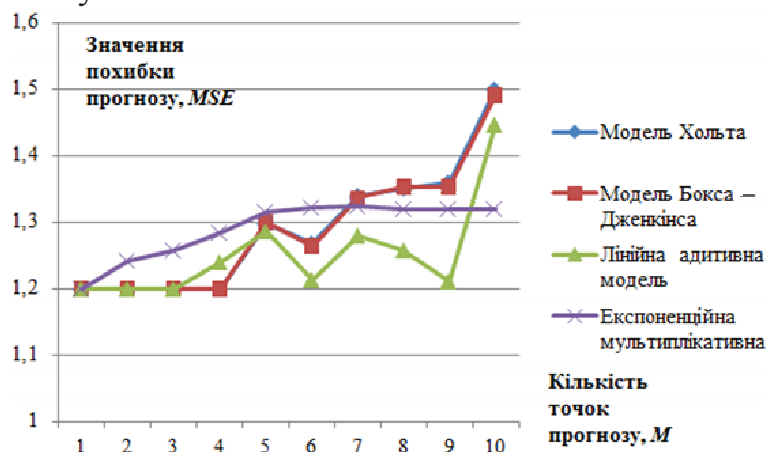
Рисунок 2 – Поверхня для моделі Бокса – Дженкінса

У табл. 4 наведено залежність середньоквадратичної похибки прогнозу MSE від кількості точок прогнозу M . Значення коефіцієнтів згладжування $\beta_1, \beta_2, \beta_3$ для кожної з моделей тотожно дорівнюють значенням у табл. 3.

Таблиця 4 – Залежність MSE від кількості точок прогнозу

Модель	MSE для $M=1$	MSE для $M=2$	MSE для $M=3$	MSE для $M=4$	MSE для $M=5$	MSE для $M=6$	MSE для $M=7$
Модель Хольта	1,2	1,2	1,2	1,2	1,297	1,268	1,339
Модель Брауна	1,4	1,42	1,44	1,445	1,517	1,41	1,46
Модель Тейла – Вейджа	1,2	1,2	1,21	1,21	1,308	1,278	1,36
Модель Бокса – Дженкінса	1,2	1,2	1,2	1,2	1,3	1,263	1,336
Лінійна адитивна модель	1,2	1,2	1,2	1,24	1,287	1,212	1,28
Лінійна мультиплікативна модель	1,21	1,24	1,25	1,274	1,304	1,306	1,306
Експоненційна адитивна модель	1,2	1,242	1,259	1,285	1,316	1,323	1,32
Експоненційна мультиплікативна модель	1,2	1,241	1,258	1,283	1,315	1,322	1,323

Нижче подано графік залежності значення MSE прогнозу від кількості точок прогнозу для двох лінійних моделей : Хольта та Бокса – Дженкінса та двох сезонних моделей: лінійної адитивної та експоненційної мультиплікативної.

Рисунок 3 – Графік залежності MSE від кількості точок прогнозу

Як видно з графіка та табл. 4, значення середньоквадратичної похибки прогнозу починає істотно збільшуватися при кількості точок прогнозу більше трьох – чотирьох.

Висновки. Адаптивні моделі прогнозування мають ряд переваг у порівнянні з багатьма іншими відомими моделями, серед яких головним є те, що вони не вимагають суттєвих математичних обчислень, є простими у реалізації, не накладають обмежень на довжину ряду та дозволяють швидко побудувати прогноз.

На досліджуваних часових рядах розроблене програмне забезпечення показало найкращі результати прогнозування при використанні моделей Хольта, Тейла – Вейджа, Бокса – Дженкінса та адитивної лінійної моделі з такими параметрами генетичного алгоритму: використання арифметичного кросовера або *SBX* – *n* кросовера за параметром $n = 5$, аутбридингу, розмірності популяції 50 особин, рівня мутації $= 0,1$ та $q = 80 \%$. В результаті проведення експериментів було встановлено, що потрібна кількість епох алгоритму залежить від розмірності ряду, якщо ряд є коротким, то задля знаходження оптимального рішення потрібно більше повторень, ніж для середніх та довгих. Окрім цього, було помічено, що алгоритм пошуку оптимальних параметрів швидше сходиться для моделей, які мають меншу кількість параметрів. Наприклад, для того щоб знайти оптимальний параметр для моделі Брауна, потрібно 10–15 епох, у той же час для того щоб налаштувати модель Бокса – Дженкінса, яка використовує три параметри, потрібно в середньому 30–40 повторень.

Також було встановлено, що з використанням реалізованого підходу якісний прогноз можна побудувати на три – чотири кроки вперед. При прогнозуванні на більшу кількість кроків вперед, середньоквадратична похибка прогнозу починає різко збільшуватися і побудований прогноз стає неадекватним.

Досліджувані моделі також мають і деякі недоліки, які, насамперед, викликані особливістю усіх адаптивних моделей – вони будують прогнозне значення із запізненням на один крок. У випадках, коли має місце різкий спад чи різкий зріст рівнів ряду, або коли знак тенденції змінюється зі зросту на спад чи зі спаду на зріст, такі моделі не можуть спрогнозувати майбутні значення правильно. Тому метою подальших досліджень є пошук методів, які були б більш гнучкими, могли аналізувати часовий ряд на присутні в ньому тенденції та шаблони і, таким чином, будували більш точний прогноз або надавали вірогідність декількох можливих майбутніх значень.

Бібліографічні посилання

1. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов: учебное пособие. Москва. 2003. 416 с.
2. Білобородько О. І., Ємельяненко Т. Г. Аналіз динамічних рядів: навчальний посібник. Дніпропетровськ. 2014. 80 с.
3. Метод групового урахування аргументів (МГУА) / URL: <http://www.mgua.irtc.org.ua/ru/index.php?page=gmdh> (дата звернення 5.11.2016).
4. Herrera F., Lozano M., Verdegay J. L. Tackling real-coded Genetic algorithms: operators and tools for the behaviour analysis // Artificial Intelligence Review. Vol. 12. No. 4. 1998. P. 265–319.

Надійшла до редколегії 05.11.16.