

# КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

## ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет комп'ютерних наук та кібернетики

Кафедра прикладної статистики

### Кваліфікаційна робота

на здобуття ступеня бакалавра

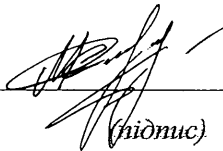
за спеціальністю 124 Системний аналіз

на тему:

### АДАПТИВНІ МОДЕЛІ КОРОТКОСТРОКОВОГО ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

Виконала студентка 4 курсу

Моцик Софія Русланівна



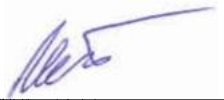
---

(підпис)

Керівник дипломної роботи

Асистент, кандидат фізико-математичних наук

Макушенко Ігор Анатолійович

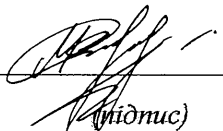


---

(підпис)

Засвідчую, що в цій роботі немає запозичень з праць інших авторів без відповідних посилань.

Студент



---

(підпис)


Роботу розглянуто й допущено до захисту на засіданні кафедри прикладної статистики

« 06 » червня 2022 р.,

протокол № 11

Завідувач кафедри

Розора І. В.



---

(підпис)

Київ - 2022

## ЗМІСТ

<b>АНОТАЦІЇ.....</b>	<b>3-4</b>
<b>ВСТУП.....</b>	<b>5</b>
<b>РОЗДІЛ 1. Найпростіші адаптивні моделі та їх властивості.....</b>	<b>7-22</b>
1.1. Часові ряди і стохастичні процеси.....	7
1.2. Експоненціальне згладжування.....	8
1.3. Початкові умови експоненційного згладжування.....	10
1.4. Вибір постійного згладжування.....	11
1.5. Реакція моделі на деякі стандартні вхідні потоки даних.....	13
1.6. Властивість оптимальності.....	15
1.7. Моделі лінійного росту.....	17
1.8. Стохастичний процес Тейла і Вейджа.....	19
<b>РОЗДІЛ 2. Розвиток моделей з постійним параметром адаптації.....</b>	<b>23-36</b>
2.1. Адаптивна модель для вивчення еволюціонуючих законів розподілу ймовірностей.....	23
2.2. Сезонні моделі.....	26
2.3. Апроксимація полінеальних трендів за допомогою багатократних згладжувань.....	31
2.4. Загальна модель Брауна.....	34
<b>РОЗДІЛ 3. Адаптивна модель прогнозування часового ряду, генерованого авторегресивною схемою з дрейфуючими коефіцієнтами.....</b>	<b>37-42</b>
3.1. Загальна схема адаптивного фільтру.....	37
3.2. Адаптація коефіцієнтів моделей авторегресії.....	39
<b>РОЗДІЛ 4. Доведення коректності алгоритму на прикладі.....</b>	<b>43-51</b>
4.1. Використання алгоритму в EXEL.....	43
4.2. Інсталювання в програмному кодї.....	47
<b>ВИСНОВКИ.....</b>	<b>52</b>
<b>ДОДАТОК А. Графік 4.1.1.....</b>	<b>53</b>
<b>ДОДАТОК Б. Графік 4.2.1.....</b>	<b>54</b>
<b>ДОДАТОК В. Графік 4.2.2.....</b>	<b>55</b>
<b>ДОДАТОК Г. Текст програми.....</b>	<b>56</b>
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....</b>	<b>58</b>

## Анотації

Дипломна робота складається і вступу, 3 розділів, висновків, списку використаних джерел. Загальний обсяг роботи становить 58 сторінки, основний текст роботи викладено на 42 сторінках.

Ключові слова: часові ряди, експоненціальне згладжування, стандарті хідні потоки даних, оптимальність, стохастичний процес Тейла і Вейджа, закони розподілу, адаптивні моделі, сезонні моделі, апроксимація, модель Брауа, адаптивний фільтр, авторегресія.

Дана робота присвячена одному з сучасних напрямків статистичного аналізу і прогнозування часових рядів. Важливість цього напрямку не визиває сумніву, так як, необхідність розв'язку відповідних задач за допомогою адаптивних методів виникає наймовірніше часто. Адаптивні методи можуть застосовуватися для прогнозування показників фондового ринку, грошових потоків, змін щоденних залишків на складах та крамницях. За допомогою цих же методів вдається описати еволюцію зміни техніко-економічних характеристик виробів і змінних параметрів хімічних процесів, вивчити поведінку показника частоти відмов обладнання в залежності від його віку. Нарешті, названі методи корисні при аналізі сезонних явищ. У ряді випадків ці методи можуть з успіхом застосовуватися для прогнозування макропоказників. Методи адаптивного прогнозування застосовуються там, де основною інформацією для прогнозу є тимчасові ряди.

Інструментом прогнозу при адаптивному методі служить модель. Первісна оцінка параметрів цієї моделі ґрунтується на даних базового (вихідного) часового ряду. На основі нових даних, одержуваних на кожному наступному кроці, відбувається коригування параметрів моделі в часі, їх адаптації до нових, безперервно мінливих умов розвитку явища. Таким чином, модель постійно «вбирає» нову інформацію і пристосовується до неї.

## **Annotations**

Thesis consists of an introduction, 3 chapters, conclusions, a list of sources used. The total volume of the work is 58 pages, the main text of the work is set out on 42 pages.

Keywords: time series, exponential smoothing, standards, data flow, optimality, Tayler and Wage stochastic process, distribution laws, adaptive models, seasonal models, approximation, Braw model, adaptive filter, autoregression.

This work is devoted to one of the modern areas of statistical analysis and forecasting of time series. There is no doubt about the importance of this area, as the need to solve relevant problems using adaptive methods is incredibly common. Adaptive methods can be used to forecast stock market performance, cash flow, changes in daily balances in warehouses and stores. Using the same methods, it is possible to describe the evolution of changes in the technical and economic characteristics of products and variable parameters of chemical processes, to study the behavior of the failure rate of equipment depending on its age. Finally, these methods are useful in the analysis of seasonal phenomena. In some cases, these methods can be successfully used to predict macro indicators. Adaptive forecasting methods are used where the main information for the forecast is time series.

The model serves as a forecasting tool for the adaptive method. The initial estimate of the parameters of this model is based on the data of the base (original) time series. Based on the new data obtained at each subsequent step, the parameters of the model are adjusted over time, their adaptation to new, continuously changing conditions of the phenomenon. Thus, the model constantly "absorbs" new information and adapts to it.

## **Вступ.**

Економічне прогнозування характеризує майбутній розвиток, виходячи з гіпотези, що основні фактори та тенденції минулого періоду зберуться на період прогнозу або що можна обґрунтувати та врахувати напрямок їх змін у аналізованій перспективі.

Надії покладаються на інерційність економічних систем. Тим часом у більшості випадків рухливість економічних явищ зростає. Найбільшу інерційність мають макроекономічні характеристики. Для параметрів, що описують процеси, що відбуваються на рівні галузей, підприємств, цехів, характерна велика залежність від місцевих умов. У сучасних умовах дослідник часто має справу з новими економічними явищами з короткими. Статистичними рядами або зі старими явищами, що зазнають докорінних змін, тому при використанні інформації для побудови моделей встає питання про наступності даних. Застарілі дані при моделюванні часто виявляються марними і навіть шкідливими. До того ж статистичний опис процесу рідко може задовольнити, тому що необхідно знати, не як розвивається процес у середньому, а як розвиватиметься його тенденція, яка існує в даний момент.

Отже, треба будувати моделі, спираючись переважно на малу кількість найсвіжіших даних. У цьому випадку альтернативою статистичного обґрунтування моделі може бути наділення її адаптивними властивостями. Спрямований розвиток економічної системи прокладає собі шлях через поодинокі явища в умовах зіткнення протидіючих сил і тому, як правило, передбачається стохастичним процесом. Для прогнозування стаціонарних процесів математичний апарат розроблено досить добре. Але економічні процеси, зазвичай, є нестаціонарними. Чим більший період прогнозу (упередження), тим більше можливостей зміни тенденцій економічного розвитку, особливо в сучасних умовах. Крім того, в досліджуваному періоді можуть статися взагалі непередбачувані, непередбачувані події, що

істотно деформують процес, що вивчається. До непередбачуваних подій відносяться такі, для передбачення яких у момент складання прогнозу немає необхідних даних, або такі, природа яких невизначена.

Основним засобом аналізу та прогнозу тимчасового ряду буде модель. Поняття модель використовується у двох значеннях: як модель тимчасового ряду, що виражає закон генерування членів ряду, та як прогнозна модель, або предиктор. Головна відмінність цих двох типів моделей у тому, що на виході моделі тимчасового ряду фактичні члени ряду, а на виході прогнозної моделі оцінки майбутніх членів низки. Теоретично властивості предиктора досліджуються в припущенні, що він застосований для отримання прогнозів процесу, що генерується моделлю, заданої аналітично.

На тимчасовий ряд впливають у різні часи різні чинники. Деякі з тих чи інших причин послаблюють свій вплив, інші впливають активніше. Таким чином, реальний процес протікає в умовах, що змінюються, що становлять його зовнішнє середовище, до якого він пристосовується, адаптується. А модель, у свою чергу, адаптується до ряду, який представляє цей процес.

Оскільки ми розглядаємо варіюючи, нестаціонарні ряди, у яких рівень, швидкість лінійного зростання та інші характеристики не залишаються постійними у часі, модель завжди перебуватиме у русі. Образно кажучи, процес адаптації моделі до ряду можна було б назвати гонкою за лідером. Очевидно, важко провести чітку грань, що відокремлює адаптивні методи прогнозування від неадаптивних. Вже прогнозування методом екстраполяції звичайних регресійних кривих містить певний елемент адаптації, коли з кожним новим отриманням фактичні дані параметри регресійних кривих перераховуються, уточнюються.

Отже, в цій роботі ми детально познайомимося з найпростішими адаптивними моделями та їх властивостями, а також розглянемо різні методи розвитку моделей з постійним параметром адаптації і все це детально опрацюємо на прикладах.

## РОЗДІЛ 1. Найпростіші адаптивні моделі та їх властивості

### 1.1. Часові ряди і стохастичні процеси

Часовий ряд - це безліч спостережень, одержуваних послідовно в часі. Якщо час змінюється дискретно, часовий ряд називається дискретним. Ми будемо розглядати тільки дискретні часові ряди, в яких спостереження робляться через фіксований інтервал часу, який приймається за одиницю рахунку. Перехід від моменту одного спостереження до моменту наступного спостереження будемо називати кроком.

Якщо значення членів тимчасового ряду точно визначені будь-якої математичної функцією, то часовий ряд називається детермінованим. Якщо ці значення можуть бути описані тільки за допомогою розподілу ймовірностей, часовий ряд називається випадковим.

Явище, що розповсюджується в часі відповідно до законів теорії ймовірностей, називається стохастичним процесом. Надалі будемо називати його просто процесом.

Аналізований відрізок тимчасового ряду може розглядатися як одна приватна реалізація (вибірка) досліджуваного стохастичного процесу, що генерується прихованим імовірнісним механізмом.

Серед випадкових процесів виділяють клас процесів, званих стаціонарними. По-значимо член тимчасової ряду, спостережний в момент часу  $t$ , через  $x_t$ . Стохастичний процес називається стаціонарним, якщо його властивості не змінюються з часом. Зокрема, він має постійне математичне очікування  $\bar{x} = M(x_t)$  (тобто, середнє значення, щодо якого він варіює), постійну дисперсію  $D(x) = M[(x_t - \bar{x})^2] = \sigma_x^2$ , визначаючи розмах його коливань щодо середнього значення, а також постійну автоковаріацію і коефіцієнти автокореляції. Коваріація між значеннями  $x_t$  і  $x_{t+k}$ , відокремленими інтервалом в  $k$  одиниць часу, називається автоковаріацією з лагом (затримкою)  $k$  і визначається як:

$$R_{xx}(k) = cov(x_t, x_{t+k}) = M[(x_t - \bar{x})(x_{t+k} - \bar{x})].$$

Для стаціонарних процесів автоковаріація залежить тільки від лага  $k$  і

$R_{xx}(0) = \sigma_x^2$ . Автокореляція з лагом  $k$  є лише нормованої автоковаріації і дорівнює:

$$\rho_k = \frac{M[(x_t - \bar{x})(x_{t+k} - \bar{x})]}{\sqrt{M[(x_t - \bar{x})^2]M[(x_{t+k} - \bar{x})^2]}} = \frac{M[(x_t - \bar{x})(x_{t+k} - \bar{x})]}{\sigma_x^2},$$

Так, як для стаціонарного процесу  $\sigma_x^2 = \text{const}$ . Таким чином,  $k$  – й коефіцієнт автокореляції  $\rho_k = \frac{R_{xx}(k)}{R_{xx}(0)}$ . Він володіє такою властивістю, що  $-1 \leq \rho_k \leq 1$ .

Для опису часових рядів використовуються математичні моделі. Уявімо, що часовий ряд  $x_t$ , що генерується деякою моделлю, можна представити у вигляді двох компонент  $x_t = \xi_t + \varepsilon_t$ , де величина  $s_t$  генерується випадковим неавтокорельованим процесом з нульовим математичним очікуванням і кінцевої (не обов'язково постійної) дисперсією, а величина  $\xi_t$  може бути генерована або детермінованою функцією, або випадковим процесом, або який-небудь їх комбінацією. Величини  $\varepsilon_t$  і  $\xi_t$  розрізняються характером впливу на значення наступних членів ряду. Мінлива  $\varepsilon_t$  впливає тільки на значення синхронного їй члена ряду, в той час як величина  $\xi_t$  певною мірою визначає значення декількох або всіх наступних членів ряду. Через величину  $\xi_t$  здійснюється взаємодія членів ряду; таким чином, в ній міститься інформація, необхідна для отримання прогнозів.

Назвемо величину  $\xi_t$  рівнем ряду в момент  $t$ , а закон еволюції рівня в часі – трендом. Таким чином, тренд може бути виражений як детермінованою, так і випадкової функціями, або їх комбінацією. Стохастичні тренди мають, наприклад, ряди з випадковим рівнем або випадковим стрибкоподібним характером росту.

## 1.2. Експоненціальне згладжування.

Припустимо, що досліджується часовий ряд  $x_t$ .

Виявлення і аналіз тенденції динамічного ряду часто виробляється за допомогою його вирівнювання або згладжування. Експоненціальне згладжування – один з

найпростіших і поширених прийомів вирівнювання ряду. В його основі лежить розрахунок експоненційних середніх.

Експоненціальне згладжування ряду здійснюється по рекурентній формулі:

$$S_t = \alpha x_t + \beta S_{t-1}$$

Де  $S_t$  – значення експоненційного середнього в момент  $t$ ;  $\alpha$  – параметр згладжування, при чому  $\alpha = \text{const}$ ,  $0 < \alpha < 1$ ;  $\beta = 1 - \alpha$ .

Цей вираз можна переписати наступним чином:

$$S_t = \alpha x_t + (1 - \alpha)S_{t-1} = S_{t-1} + \alpha(x_t - S_{t-1}).$$

Експоненціальна середня на момент  $t$  тут виражена як експоненціальна середня попереднього моменту плюс частка  $\alpha$  різниці поточного спостереження і експоненційної середньої минулого моменту.

Якщо послідовно використовувати перше рекурентне співвідношення, то експонентну середню  $S_t$  можна виразити через значення часового ряду  $x_t$ :

$$\begin{aligned} S_t &= \alpha x_t + \beta S_{t-1} = \alpha x_t + \beta x_{t-1} + \beta^2 S_{t-2} = \dots = \\ &= \alpha x_t + \alpha \beta x_{t-1} + \alpha \beta^2 x_{t-2} + \dots + \alpha \beta^i x_{t-i} + \dots + \beta^N S_0 = \\ &= \alpha \sum_{i=0}^{N-1} \beta^i x_{t-i} + \beta^N S_0, \end{aligned}$$

Де  $N$  – кількість членів ряду;  $S_0$  – деяка величина, яка характеризує початкові умови для першого використання формули рівняння один при  $t = 1$ .

Так як,  $\beta < 1$ , то при  $N \rightarrow \infty$   $\beta^N \rightarrow 0$ , а сума коефіцієнтів  $\alpha \sum_{i=0}^{N-1} \beta^i \rightarrow 1$ . Тоді

$$S_t = \alpha \sum_{i=0}^{\infty} \beta^i x_{t-i}.$$

Таким чином, величина  $S_t$  виявляється ваговою сумою всіх членів ряду. Причому ваги падають експоненціально залежно від давності («віку») спостереження. Це і пояснює, чому величина  $S_t$  названа експоненційною середньою.

Розглянемо, ряд, генерований моделлю:  $x_t = a_1 + \varepsilon_t$ ,

Де  $a_1 = \text{const}$ ;  $\varepsilon_t$  - випадкові неавтокорельованні відхилення, або шум, із середнім значенням 0 і дисперсією  $\sigma^2$ .

Застосуємо до нього процедуру експоненціального згладжування з рівняння один. Тоді

$$S_t = \alpha \sum_{i=0}^{\infty} \beta^i x_{t-i} = \alpha \sum_{i=0}^{\infty} \beta^i (a_1 + \varepsilon_t) = a_1 + \alpha \sum_{i=0}^{\infty} \beta^i \varepsilon_{t-i}.$$

Знайдемо математичне сподівання

$$M(S_t) = M(x_t) = a_1$$

І дисперсію

$$D(S_t) = M[(S_t - a_1)^2] = M\left[\left(\alpha \sum_{i=0}^{\infty} \beta^i \varepsilon_{t-i}\right)^2\right] = \alpha^2 \sum_{i=0}^{\infty} \beta^{2i} \sigma^2 = \frac{\alpha}{2 - \alpha} \sigma^2.$$

Так, як  $0 < \alpha < 1$ ,  $D(S_t) < D(x_t) = \sigma^2$ . Таким чином, експоненціальна середня  $S_t$  має те ж математичне очікування, що і ряд  $x$ , але меншу дисперсію. Як видно з попереднього рівняння, при високому значенні  $\alpha$  дисперсія експоненційної середньої незначно відрізняється від дисперсії ряду  $x$ . Чим менше  $\alpha$ , тим більшою мірою скорочується дисперсія експоненційної середньої.

Отже, експоненціальне згладжування можна уявити як фільтр, на вхід якого у вигляді потоку послідовно надходять члени вихідного ряду, а на виході формуються поточні значення експоненційної середньої. І чим менше  $\alpha$ , тим більшою мірою фільтруються, по переважна коливання вихідного ряду.

### 1.3. Початкові умови експоненційного згладжування.

Експоненціальне вирівнювання завжди вимагає попереднього значення експоненційної середньої. Коли процес тільки починається, повинна бути деяка величина  $S_0$ . Яка може бути використана в якості значення, що передує  $S_1$ .

Якщо є минулі дані до моменту початку вирівнювання, то в якості початкового значення  $S_0$  можна використовувати арифметичну середню всіх наявних точок

або якийсь їх частини. Коли для такого оцінювання  $S_0$  немає даних, потрібно проорокування початкового рівня ряду.

Передбачення може бути зроблено виходячи з апріорних знань про процес або на основі його аналогії з іншими процесами. Після  $k$  кроків вага, якої надає початкового значення, дорівнює  $(1 - \alpha)^k$ . Якщо є впевненість в справедливості початкового значення  $S_0$ , то можна коефіцієнт  $\alpha$  взяти малим. Якщо такої впевненості немає, то параметру  $\alpha$  слід дати велике значення, з таким розрахунком, щоб вплив початкового значення швидко зменшилася. Однак велике значення  $\alpha$ , яке впливає з останнього рівняння в попередньому розділі, може з'явитися причиною великої дисперсії коливань  $S_t$ . Якщо необхідно придушення цих коливань, то після достатнього віддалення від початкового моменту часу величину  $\alpha$  можна відняти.

#### 1.4. Вибір постійного згладжування

Вибору величини постійної згладжування слід приділяти особливу увагу. Пошуки повинні бути спрямовані на відшукання підстав для вибору найкращого значення. Потрібно враховувати умови, при яких ця величина повинна приймати значення, близькі то одному крайнього значенням, то іншому. Неважко помітити, що при  $\alpha = 0$   $S_t = S_0$  представляє випадок абсолютної фільтрації і повної відсутності адаптації, а при  $\alpha = 1$  приходимо до так званої найвної моделі  $\hat{x}_t(t) = S_t = x_t$  відповідно до якої прогноз на будь-який термін дорівнює поточному фактичним значенням ряду. На практиці ця модель через за простоти користується особливою популярністю.

В попередніх розділах ми вже відзначали, що постійна згладжування характеризує швидкість реакції моделі  $\hat{x}_t(t) = S_t$  на зміни рівня процесу, але одночасно визначає і здатність системи згладжувати випадкові відхилення. Тому величиною  $\alpha$  слід давати ту чи іншу проміжне значення між 0 і 1 в залежності від конкретних властивостей динамічного ряду.

В якості задовільного компромісу Р. Браун рекомендує брати  $\alpha$  в межах від 0,1 до 0,3. Ця рекомендація не критично повторена в ряді робіт. Тим часом показано, що навіть при прогнозуванні ряду, використаного Брауном для ілюстрації, найкращі результати виходять при  $\alpha = 0,9$ . Наш досвід роботи з економічними рядами показує, що найбільша точність прогнозування може бути досягнута при будь-яких допустимих значеннях  $\alpha$ . Однак, як правило, якщо в результаті випробувань виявлено, що оптимальне значення константи  $\alpha$  близько до 1, слід перевірити законність вибору моделі даного типу. Часто до більших значень  $\alpha$  призводить наявність в досліджуваному ряді яскраво виражених тенденцій або сезонних коливань. У цьому випадку для отримання ефективних прогнозів потрібна інша модель.

Ясно, що оптимальне значення  $\alpha$  в загальному випадку має залежати від терміну прогнозування  $t$ . Для кон'юнктурних прогнозів в більшій мірі повинна враховуватися свіжа інформація. При збільшенні періоду попередження  $t$  пізніша інформація, яка відображає останню кон'юнктуру, повинна, очевидно, мати в декілька раз меншу вагу, ніж в разі малих  $t$ . Для того щоб згладити кон'юнктурні коливання, слід в більшій мірі враховувати інформацію за минулі періоди часу. Для проведення такого аналізу вводять поняття середнього віку даних.

Вік поточного спостереження дорівнює 0, вік попереднього спостереження дорівнює 1 і т. д. Середній вік - це сума зважених віку даних, використаних для підрахунку згладженої величини. Причому віку мають ті ж ваги, що і відповідна інформація. При експоненційному вирівнюванні вага, що дається точці з віком  $k$ , дорівнює  $\alpha\beta^k$ , де  $\alpha - 1$  і середній вік інформації дорівнює:

$$k = 0 \cdot \alpha + 1 \cdot \alpha\beta + 2 \cdot \alpha\beta^2 + \dots = \alpha \sum_{k=0}^{\infty} k\beta^k = \frac{\beta}{\alpha}.$$

Таким чином, чим менше  $\alpha$ , тим більший середній вік інформації.

## 1.5. Реакція моделі на деякі стандартні вхідні потоки даних

Розглянемо, які реакції моделі експоненціального згладжування на деякі стандартні вхідні потоки, які містять типові порушення стаціонарності, і як позначається на цих реакціях величина постійної згладжування. Такі вхідні потоки носять абстрактний характер, проте проводиться з їх допомогою аналіз дозволяє глибше вивчити адаптивні властивості моделі.

*Реакція на імпульс.*

Першим і найбільш важливим тестом є імпульс (Дельта-функція Кронекера):

$$x_t = \delta(t) \text{ або } x_0 = 1; \quad x_t = 0 \text{ при } t \neq 0.$$

Поодинокі імпульси являють собою події, викликані сторонніми для досліджуваного явища причинами. В економіці імпульс відображає разове явище, яке має місце лише в даний момент часу. У загальному випадку вимога до прогнозуючої системі полягає в тому, щоб її реакція на імпульс була якомога слабкіше, тому що дія імпульсу короткочасно і не буде проявлятися в майбутньому. Ця реакція характеризує фільтруючі властивості системи.

Реакція на імпульс є функцією часу  $h_t$ . Вона описує вихід системи через  $t$  одиниць часу після надходження одиничного імпульсу на вхід. Будь-дискретний часовий ряд можна розглядати як серію імпульсів відповідної амплітуди. Якщо реакція лінійно системи на одиничний імпульс є  $h_t$  то її реакція на послідовність імпульсів  $x$  в момент  $t$  визначається як результат додавання реакцій на кожен імпульс, тобто як сума:

$$y_t = \sum_{i=0}^{\infty} h_i x_{t-i}.$$

Реакція моделі експоненціального згладжування, як відомо, є лінійною функцією членів тимчасового ряду і має вигляд:

$$S_t = \sum_{i=0}^{\infty} \alpha \beta^i x_{t-i}.$$

З цих двох рівнянь слідує, що  $h_i = \alpha\beta^i$ . Таким чином, можна говорити або про ваги, з якими зважуються члени ряду, або про реакцію на одиничний імпульс. Будь-яка дискретна система, що виражається за допомогою лінійних рівнянь в кінцевих різницях з постійними коефіцієнтами, може бути описана за допомогою її реакції на одиничний імпульс.

*Реакція на ступеневу зміну.*

При наявності істотних довгострокових змін структури ряду необхідно домагатися, щоб модель враховувала їх якомога швидше. Розглянемо насамперед ступеневу зміну рівня вхідного потоку. Запишемо його аналітично:

$$x_t = a_1 \text{ для } t < t_1; \quad x_t = a_2 \text{ для } t \geq t_1 \quad a_1 \neq a_2.$$

В економічних дослідженнях ступеневу зміну може відображати різка зміна рівня виробництва, попиту або споживання будь-якого товару в результаті повороту моди, впровадження значного наукового відкриття, зміни політичної або зовнішньоекономічної ситуації.

Одиничний ступінчастий стрибок є наступною функцією:

$$x_t = 0 \text{ для } t < 0; \quad x_t = 1 \text{ для } t \geq 0.$$

Оскільки розглянута процедура згладжування є лінійною, то можна визначити реакцію на такого роду зміни за допомогою реакції на одиничний імпульс.

Реакція моделі буде:

$$S_t = \sum_{i=0}^{\infty} h_i x_{t-i} = \sum_{i=0}^t \alpha \beta^i = 1 - \beta^{t+1}.$$

З кроком зростання  $t$  член  $\beta^{t+1}$  стає незначним і величина  $S_t$  наближається до рівня  $x_t$ .

*Реакція на лінійну і квадратичну функції.*

Тимчасові економічні ряди часто мають тенденцію лінійного або параболічного росту. Можна показати, що при лінійно-наростаючому вхідному потоці  $a_1 + a_2 t$

експоненціальна середня буде постійно відставати від часового ряду і що це відставання в кінці кінців прагне до величини  $\frac{\beta}{\alpha} a_2$  тобто, чим більше  $\alpha$ , тим менше відставання.

При квадратичному вхідному потоці типу  $a_1 + a_2 t + a_3 t^2$  експоненційна середня відстає від часового ряду все більше і більше. Найменша відставання буде при  $\alpha \approx 1$ .

При прогнозуванні в цих обох випадках не можна від моделі експоненційної середньої очікувати хороших прогнозів. Для таких часових рядів потрібні інші моделі.

*Реакція на синусоїдальну хвилю.*

У багатьох випадках тимчасові ряди по своїй природі є періодичними. Браун показав, що реакція експоненційної середньої на вхідний потік  $x_t = \sin \frac{2\pi t}{12}$  буде:

$$S_t = \frac{\alpha \beta^{t+1}}{2(1 - \sqrt{3}\beta + \beta^2)} + \frac{\alpha}{\sqrt{1 - \sqrt{3}\beta + \beta^2}} \sin \frac{2\pi(t - \varphi)}{12},$$

Де фазовий кут  $\varphi$  визначається співвідношенням:

$$\operatorname{tg} \varphi = \frac{\beta}{\sqrt{3}\beta - 2}.$$

Перший доданок прямує до нуля з плином часу, і в результаті експоненційна середня буде також змінюватися по синусоїді того ж самого періоду, але з амплітудою і кутовим зсувом, залежними від  $\alpha$ .

## 1.6. Властивість оптимальності

Головне достоїнство прогновної моделі, заснованої на експоненційній середній, полягає в тому, що вона здатна послідовно адаптуватися до нового рівня процесу без значного реагування на випадкові відхилення. Однак доцільно визначити статистичні властивості таких часових рядів, за якими цей метод прогнозування

працює особливо добре. В результаті можна буде краще судити про сферу застосування цього методу або модифікувати його в тому випадку, коли необхідні властивості у часового ряду відсутні.

Першим хто досліджував дане питання був Д. Мат. Він взяв часовий ряд  $x$ , згенерований математичною моделлю:

$$x_t = \xi_t + \varepsilon_t ,$$

Де  $\varepsilon_t$  – випадкові незалежні відхилення з середнім значенням 0 і дисперсією  $\sigma_\varepsilon^2$ ;  $\xi_t$  – величина, яка отримує на кожному кроці деякий приріст  $u_t$ , тобто:

$$\xi_t = \xi_{t-1} + u_t = \sum_{i=1}^t u_i ,$$

Де величини  $u_i$  незалежні, мають середнє значення 0 і дисперсію  $\sigma_u^2$ .

Значення величин  $\varepsilon$  і  $u$  поки передбачаються незалежними. Такий часовий ряд можна розглядати як випадкове рух рівня процесу  $\xi_t$ , на яке накладено шум  $\varepsilon_t$ .

Д. Мат поставив задачу відшукати оптимальні ваги  $\omega_k$  в предикативі виду:

$$\widehat{x}_1(t) = \sum_{k=0}^{\infty} \omega_k x_{t-k} ,$$

які мінімізують дисперсію помилки прогнозу ряду зазначеного вище. Він прийшов до висновку, що оптимальними вагами в цій моделі будуть:

$$\omega_k = (1 - \lambda)\lambda^k , k = 0, 1, 2, \dots ;$$

$$\lambda = 1 + \frac{1}{2} \frac{\sigma_u^2}{\sigma_\varepsilon^2} - \frac{\sigma_u}{\sigma_\varepsilon} \sqrt{1 + \frac{1}{4} \frac{\sigma_u^2}{\sigma_\varepsilon^2}} < 1 .$$

Ваги мають ту ж саму форму, що й у моделі експоненціальної середньої, яка визначається:

$$\widehat{x}_1(t) = \alpha \sum_{k=0}^{\infty} (1 - \alpha)^k x_{t-k} ,$$

якщо покласти  $\alpha = 1 - \lambda$ . Отже, модель експоненційної середньої в даному випадку є оптимальною при:

$$\alpha = \frac{\sigma_u}{\sigma_\varepsilon} \sqrt{1 + \frac{1}{4} \frac{\sigma_u^2}{\sigma_\varepsilon^2} - \frac{1}{2} \frac{\sigma_u^2}{\sigma_\varepsilon^2}}.$$

Якщо дисперсія збільшень рівня  $\sigma_u^2$  мала в порівнянні з дисперсією шуму про  $\sigma_\varepsilon^2$ , то  $\alpha$  буде близька до 0. Прогнози в такому випадку мало залежать від нової інформації. Низьке значення  $\alpha$  забезпечує хорошу фільтрацію шуму. І навпаки, якщо  $\sigma_u^2$  велика в порівнянні з дисперсією шуму, то  $\alpha$  буде близька до 1, так що вага нової інформації зростає.

У тому випадку, якщо  $\varepsilon$  і  $u$  корельовані і

$$\begin{aligned} M(\varepsilon_t u_t) &= \sigma_{\varepsilon u}; \\ M(\varepsilon_t u_{t'}) &= 0 \quad t \neq t', \end{aligned}$$

необхідно лише у виразі для  $\alpha$  змінити співвідношення

$$\frac{\sigma_u^2}{\sigma_\varepsilon^2} \quad \text{на} \quad \frac{\sigma_u^2}{\sigma_\varepsilon^2 + \sigma_{\varepsilon u}}.$$

### 1.7. Моделі лінійного росту

В 2 розділі «Реакція моделі на деякі стандартні вхідні потоки даних», показано, що експоненційна середня призводить до зміщених прогнозів, тобто дає систематичну помилку, коли часовий ряд має тенденцію лінійного росту. Для цього випадку розроблено кілька варіантів адаптивних моделей, які також використовують процедуру експоненціального згладжування. В основі моделей лежить гіпотеза про те, що прогноз може бути отриманий за рівнянням

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

Де  $\hat{a}_{1,t}$ ,  $\hat{a}_{2,t}$  - поточні оцінки коефіцієнтів адаптуюся ного полінома першого порядку.

Однією з перших моделей цього типу була двопараметрична модель Ч. Хольта, в якій оцінка коефіцієнтів проводиться таким чином:

$$\widehat{a}_{1,t} = \alpha_1 x_t + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1});$$

$$\widehat{a}_{2,t} = \alpha_2(\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_2)\widehat{a}_{2,t-1},$$

де  $\alpha_1$  і  $\alpha_2$  – параметри експоненційного згладження ( $0 < \alpha_1; \alpha_2 < 1$ ), які ми також будемо називати параметрами адаптації.

Окремим випадком моделі Хольта являється модель лінійного росту Брауна:

$$\widehat{a}_{1,t} = \widehat{a}_{1,t-1} + \widehat{a}_{2,t-1} + (1 - \beta^2)e_t;$$

$$\widehat{a}_{2,t} = \widehat{a}_{2,t-1} + (1 - \beta^2)e_t,$$

де параметр  $\beta$  – коефіцієнт дисконтування, який характеризує знецінення даних спостереження за одиницю часу,  $0 < \beta < 1$ ;

$e_t = x_t - \widehat{x}_1(t-1)$  – помилка прогнозу.

Якщо модель Хольта вдосконалити шляхом включення різниці помилок, то отримаємо повну три-параметричну модель прогнозування Дж. Боксу і Г.Дженкінса:

$$\widehat{x}_t(t) = \widehat{a}_{1,t} + \tau \widehat{a}_{2,t};$$

$$\widehat{a}_{1,t} = \alpha_1 x_t + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1}) + \alpha_3(e_t - e_{t-1});$$

$$\widehat{a}_{2,t} = \alpha_2(\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_2)\widehat{a}_{2,t-1},$$

де  $\alpha_1, \alpha_2, \alpha_3$  – являються параметром моделі,  $0 < \alpha_1, \alpha_2, \alpha_3 < 1$ ;

$e_t = x_t - \widehat{x}_1(t-1)$  – помилка прогнозу.

На основі практичних випробувань моделі на багатьох економічних рядах Бокс і Дженкінс дійшли висновку, що включення до моделі різниці помилок не є необхідним. Коефіцієнт  $\alpha_3$  завжди опинявся близьким до нуля. П. Харрісон прийшов до такого ж висновку. Це пояснюється стохастичним характером даних, і зокрема, тим, що кореляція помилок в подібних випадках нестійка.

Харрісон провів емпіричне порівняння одно-параметричної моделі Брауна з багато-параметричними моделями. Багато-параметричні моделі в жодному випадку не дали помітної переваги. Тому на практиці для прогнозування рядів з лінійної тенденцією краще використовувати більш просту модель Брауна.

## 1.8. Стохастичний процес Тейла і Вейджа

Г. Тейл і С. Вейдж аналогічно тому, як це зробив Д. Мат при вивченні експоненційної середньої, з метою подальшого вивчення властивостей адаптивних моделей запропонували застосувати двух-параметричний предиктор Хольта для прогнозування деякого імовірнісного процесу, що характеризується стохастичним трендом. Вони вивели вираження для визначення оптимальних параметрів адаптації, які мінімізують середній квадрат помилки прогнозування.

Процес Тейл-Вейджа аналітично записується як:

$$\begin{aligned}x_t &= a_{1t} + \varepsilon_t; \\a_{1,t} &= a_{1,t-1} + a_{2,t}; \\a_{2,t} &= a_{2,t-1} + v_t,\end{aligned}$$

де  $a_{1,t}$  – значення рівню дослідженого часового ряду  $x_t$  в момент  $t$  ;

$a_{2,t}$  – приріст рівня від моменту  $t - 1$  до моменту  $t$  ;

$\varepsilon_t, v_t$  - тимчасові послідовності з нульовим математичним очікуванням, постійними дисперсіями і відсутністю коваріації, тобто:

$$\begin{aligned}M(\varepsilon_t) &= M(v_t) = 0; \\M(\varepsilon_t \varepsilon_{t'}) &= \begin{cases} \sigma_\varepsilon^2 & t = t'; \\ 0 & t \neq t'; \end{cases} \\M(v_t v_{t'}) &= \begin{cases} \sigma_v^2 & t = t'; \\ 0 & t \neq t'; \end{cases}\end{aligned}$$

$$M(v_t \varepsilon_{t'}) = 0 \quad \text{для будь-якої пари } (t, t').$$

Часовий ряд  $x_t$  не є стаціонарним і не має чітко визначеної автоковаріаційної функції. Однак М. Нерлі і С. Вейдж показали, що з рівнянь на початку цього розділу слідує стаціонарність других різниць процесу  $x_t$ , які ми позначимо через  $y_t$ :

$$y_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = \nabla^2 x_t = \nabla^2 a_{1,t} + \nabla^2 \varepsilon_t = v_t + \nabla^2 \varepsilon_t,$$

де  $\nabla$  - різницевий оператор,  $\nabla x_t = x_t - x_{t-1}$ ,  $\nabla^2 x_t = \nabla(\nabla x_t)$ .

Другі різниці мають цілком певну автоковаріаційну функцію:

$$\text{cov}_{yy}(k) = M(y_k y_{t-k}) = \begin{cases} (g^2 + 6)\sigma_\varepsilon^2 & k = 0 \\ -4\sigma_\varepsilon^2 & |k| = 1 \\ \sigma_\varepsilon^2 & |k| = 2 \end{cases}$$

У всіх інших випадках буде 0, де

$$g^2 = \frac{\sigma_v^2}{\sigma_\varepsilon^2}.$$

Ці властивості можуть бути використані для вирішення питання про можливість адекватного уявлення тимчасового ряду процесом Тейл-Вейджа. При цьому не слід випускати з уваги ту обставину, що оцінки автоковаріаційні функції є досить грубими і корельованими і точні рівності нашої системи на практиці будуть виконуватися лише наближено.

Схема складання прогнозу відповідно до моделі Ч. Хольта виглядає наступним чином:

$$\begin{aligned} \widehat{a}_{1,t} &= \alpha_1 x_t + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1}); \\ \widehat{a}_{2,t} &= \alpha_2(\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_2)\widehat{a}_{2,t-1}; \\ \widehat{x}_t(t) &= \widehat{a}_{1,t} + \tau \widehat{a}_{2,t} \\ 0 &< \alpha_1, \alpha_2 < 1. \end{aligned}$$

Якщо помилку прогнозу, зробленого в момент  $t$  на 1 крок вперед, позначити через  $e_1(t)$ , то рівняння адаптації, які ми записали вище можна записати у вигляді:

$$\begin{aligned} \widehat{a}_{1,t} &= \widehat{a}_{1,t-1} + \widehat{a}_{2,t-1} + \alpha_1 e_1(t-1); \\ \widehat{a}_{2,t} &= \widehat{a}_{2,t-1} + \gamma e_1(t-1); \\ \gamma &= \alpha_1 \alpha_2. \end{aligned}$$

Помилка прогнозу:

$$\begin{aligned} e_1(t) &= x_{t+1} - \widehat{a}_{1,t} - \widehat{a}_{2,t} = (a_{1,t} + a_{2,t} + v_{i+1} + \varepsilon_{i+1}) - \widehat{a}_{1,t} - \widehat{a}_{2,t} = \\ &= (a_{1,t} - \widehat{a}_{1,t}) + (a_{2,t} - \widehat{a}_{2,t}) + v_{i+1} + \varepsilon_{i+1}. \end{aligned}$$

Отже, помилка прогнозу є сумою трьох компонент: помилки оцінки рівня процесу в момент  $t$ , помилки оцінки приросту рівня в момент  $t$  і комбінації випадкових компонент  $v$  і  $\varepsilon$  в момент  $t + 1$ .

Очевидно, що визначення оптимальних  $\alpha_1$  і  $\gamma$  еквівалентно визначенню оптимальних  $\alpha_1$  і  $\alpha_2$ . Оптимум зазвичай

відшукується шляхом мінімізації середнього квадрата помилки прогнозу. Але коли розглядаються нестационарні часові ряди, то в загальному випадку не очевидно, що середній квадрат помилок прогнозування адаптивним методом є величиною стійкою, яка може бути мінімізована. Використовуючи співвідношення:

$$\widehat{y_{t+1}} = \widehat{x_{t+1}} - 2x_t + x_{t-1},$$

Нерлі і Вейдж показали, що проблема прогнозування  $x_{t+1}$  еквівалентна задачі прогнозування другий різниці  $y_{t-1}$  і що при обмеженнях, накладених на параметри адаптації  $\alpha_1$  і  $\alpha_2$ , помилка прогнозу є лінійною комбінацією поточного і минулих значень стаціонарного ряду  $y_t$ :

$$e_1(t) = y_{t+1} - \sum_{k=0}^{\infty} \omega_k y_{t-k},$$

де  $\{\omega\}$  – збіжний ряд вагів.

Це означає, що помилки прогнозу стаціонарні і їх середній квадрат цілком певний.

В результаті мінімізації дисперсії помилки прогнозу на 1 крок вперед  $D_e(1)$  Тейл і Вейдж отримали наступні результати:

$$\alpha_1 = \frac{2h}{1+h}; \quad \alpha_2 = h; \quad \gamma = \frac{2h^2}{1+h};$$

$$h = \sqrt{\left(-\frac{1}{8}g^2 + \frac{1}{2}g \sqrt{1 + \frac{1}{16}g^2}\right)}; \quad g^2 = \frac{\sigma_v^2}{\sigma_\varepsilon^2};$$

$$D_e(1) = \frac{1+h}{1-h} \sigma_\varepsilon^2.$$

Груба оцінка співвідношення дисперсій  $g^2$  може бути отримана з системи по підрахованим автоковаріаціям процесу  $y_t$ . Уточнення  $g^2$  проводиться експериментально методом проб на наявному відрізку ряду. Даючи значення в околицях грубої оцінки, знаходять  $g^2$ , мінімізує дисперсію помилки  $D_e(1)$ .

Нерлі і Вейдж провели теоретичний аналіз чутливості дисперсії помилки прогнозу  $D_e(1)$  до помилки у визначенні  $g^2$ . Виявилось, що процентна зміна дисперсії  $D_e(1)$  пропорційно квадрату відносної помилки оцінки  $g^2$  з коефіцієнтом пропорційності:

$$h(1 - h)^2(6 + 8h + \frac{3h^2}{[4(2 - h^2)^3]}).$$

## РОЗДІЛ 2. Розвиток моделей з постійним параметром адаптації

У цьому розділі розглядається розвиток простішої адаптивної моделі експоненціального типу по декількох напрямленнях. Одні модифікації дозволяють застосовувати її для вивчення змінюються в часі законів розподілу ймовірностей, інші - адекватно описувати часові ряди з періодичними сезонними коливаннями, треті - апроксимувати тенденції ряду за допомогою поліномів з адаптивними коефіцієнтами

### 2.1. Адаптивна модель для вивчення еволюціонуючих законів розподілу ймовірностей

У деяких задачах потрібно визначити форму закону розподілу ймовірностей будь-якої випадкової змінної, в той час як стохастичний процес, який вона представляє, зазнає деякі зміни. В цьому випадку потрібно знайти спосіб вивчення еволюціонують законів розподілу ймовірностей.

Будемо розглядати повну систему  $n$  несумісних подій, визначених на числової осі з допомогою  $n + 1$  кордону:

$$X_0 < X_1 < X_2 < \dots < X_n .$$

Якщо досліджуються обсяги замовлень покупців на який-небудь продукт, то можуть бути, наприклад, визначимо три події:

- 1) обсяг замовлення менше 5 штук;
- 2) обсяг замовлення не менше 5 штук, але менше 20;
- 3) обсяг замовлення не менше 20 штук.

Незалежно від того, в якому обсязі продукт був замовлений, ми можемо розглядати це замовлення як здійснення одного і тільки одного з вказаних подій. Подія, пов'язане зі спостереженням  $x_t$ , відповідає номеру інтервалу, в якому виявляється спостереження. Будь-яке можливе спостереження повинно або бути рівним одній з кордонів, або опинитися між двома сусідніми межами, тобто є тільки одне  $k$ , таке, що  $X_{k-1} < x_t \leq X_k$ . Тому ми пов'язуємо зі спостереженням  $x_t$  подія

$k$ . Це означає, що перша межа подій  $X_0$  повинна бути менше, ніж будь-яке спостереження, яке може мати місце, і остання межа  $X_n$  повинна бути більше, ніж будь-яке можливе спостереження. Так як не можна бути абсолютно впевненим, що спостережувані величини будуть обмежені, то можливі дві альтернативи конструювання системи. Одна полягає в тому, щоб покласти  $X_0 = -\infty$  і  $X_n = +\infty$ .

Інший шлях - покласти  $X_0$  настільки малим, а  $X_n$  настільки великим кінцевим числом, щоб можна було очікувати, що реальні спостереження будуть перебувати в цих межах. Тоді для випадку, коли з'являється спостережене значення, що виходить за встановлені межі, слід забезпечити вироблення особливого сигналу для втручання дослідника, який повинен проаналізувати несподівані спостереження, перш ніж обробляти їх автоматично.

Розглянемо простий метод оцінки ймовірностей  $P_k(t)$  про настання різних подій:

$$X_{k-1} < x_t \leq X_k,$$

запропонований Р. Г. Брауном. Так як розглядається повна система  $n$  несумісних подій, то ступенів свободи буде  $n - 1$ , бо сума ймовірностей повинна дорівнювати одиниці, тобто. Не всі  $n$  значень  $P_k$  незалежні.

Аналіз минулих даних або судження про майбутнє робить можливим встановити межі подій і зробити початкові, хоча б грубі, оцінки ймовірностей різних подій  $P_k(0)$ ,  $k = 1, 2, \dots, n$ . Нехай спостереження  $x_t$  в момент  $t$  свідчить про настання події  $k$ . Побудуємо  $n$ -мірний вектор-стовпець  $U(t)$ , який має  $n - 1$  нульову компоненту, а  $k$ -та компонента дорівнює одиниці.

Попередні оцінки  $n$  ймовірностей можна розглядати як  $n$ -компонентний вектор-стовпець  $\hat{P}(t - 1)$ . Процес перегляду цих оцінок з урахуванням поточної інформації є експоненціальним згладжуванням за правилом

$$\hat{P}(t) = \alpha U(t) + (1 - \alpha) \hat{P}(t - 1), \quad 0 < \alpha < 1.$$

Кожна компонента вектора модифікується простим експоненціальним згладжуванням нуля або одиниці. Наприклад, якщо досліджується система з 3 подій і спостереження означає наступ другого, то

$$\begin{aligned}\widehat{P}_1(t) &= \alpha \cdot 0 + (1 - \alpha)\widehat{P}_1(t - 1); \\ \widehat{P}_2(t) &= \alpha \cdot 1 + (1 - \alpha)\widehat{P}_2(t - 1); \\ \widehat{P}_3(t) &= \alpha \cdot 0 + (1 - \alpha)\widehat{P}_3(t - 1).\end{aligned}$$

Так як  $\widehat{P}(t - 1)$  є вектором ймовірностей, то всі його компоненти повинні бути невід'ємні і їх сума повинна бути точно дорівнює 1.

Розглянутий процес векторного згладжування не може зробити компоненту негативною, і сума підсумкових компонент та ж, що сума їхніх колишніх значенні. Отже, якщо  $\widehat{P}(t - 1)$  є імовірнісним вектором, то їм і є  $\widehat{P}(t)$ .

Розглянемо подію  $i$ . Якщо закон розподілу спостережених значень  $x_t$  не змінюється, то математичне очікування значення  $i$ -ї компоненти вектора  $U(t)$ , що підлягають згладжуванню, точно так само дійсної ймовірності  $P_i$  настання події  $i$  і математичне очікування оцінки одно дійсної ймовірності

$$M[\widehat{P}_t(t)] = P_t.$$

Ймовірність того, що доведеться згладжувати одиницю, дорівнює  $P_i$ , а ймовірність того, що будемо згладжувати нуль, дорівнює  $1 - P_i$ . Легко підрахувати, що  $i$  – та компонента вектора  $U(t)$  характеризується дисперсією  $P_t(1 - P_t)$ .

Виражаючи дисперсію результату експоненціального згладжування через дисперсію на вході, отримаємо дисперсію оцінки ймовірності настання  $i$  – ї події:

$$\sigma_t^2 = \frac{\alpha}{2 - \alpha} P_t(1 - P_t),$$

де  $\alpha$  – стала згладжування.

Таким чином, є два способи конструювання системи, які найбільш кращі. Межі подію доцільно вибирати так, щоб  $P_k$  була дуже великою (близькою до 1) або

дуже маленькою (майже 0), це забезпечить малу дисперсію оцінок компонент вектора ймовірностей.

## 2.2. Сезонні моделі

В економіці багато явищ характеризуються періодично повторюваними сезонними ефектами. Відповідно тимчасові ряди, їх відображають, містять періодичні сезонні коливання. Ці ряди і їх коливання можна уявити як генеруються моделями двох основних типів: моделями з мультиплікативними і з адитивними коефіцієнтами сезонності.

Моделі першого типу мають вигляд:

$$\begin{aligned}x_t &= \xi_t + \varepsilon_t; \\ \xi_t &= a_{1,t} f_t ,\end{aligned}$$

де динаміка величини  $a_{1,t}$  характеризує тенденцію розвитку процесу;

$f_t, f_{t-1}, \dots, f_{t-l+1}$  – коефіцієнти сезонності;

$l$  – кількість фаз в повному сезонному циклі;

$\varepsilon_t$  – неавтокореляційний шум з нульовим математичним сподіванням.

Моделі другого типу мають вигляд:

$$\begin{aligned}x_t &= \xi_t + \varepsilon_t; \\ \xi_t &= a_{1,t} + g_t ,\end{aligned}$$

де величина  $a_{1,t}$  описує тенденцію розвитку процесу;

$g_t, g_{t-1}, \dots, g_{t-l+1}$  – адитивні коефіцієнти сезонності;

$l$  – кількість фаз в повному сезонному циклі;

$\varepsilon_t$  – неавтокореляційний шум з нульовим математичним сподіванням.

*Прогнозування з коефіцієнтами сезонності*

Модель має вигляд:

$$\begin{aligned}\widehat{a}_{1,t} &= \alpha_1 \frac{x_t}{\widehat{f}_{t-l}} + (1 - \alpha_1) \widehat{a}_{1,t-1} & 0 < \alpha_1 < 1; \\ \widehat{f}_t &= \alpha_2 \frac{x_t}{\widehat{a}_{1,t}} + (1 - \alpha_2) \widehat{f}_{t-l} & 0 < \alpha_2 < 1.\end{aligned}$$

Як бачимо  $\widehat{a}_{1,t}$  являється зваженою сумою даної оцінки  $\frac{x_t}{\widehat{f}_{t-l}}$ , отримана під час очищення від сезонних коливань фактичних даних  $x_t$  і попередньої оцінки  $\widehat{a}_{1,t-1}$ . В якості коефіцієнта сезонності  $\widehat{f}_t$  береться його найбільш пізня оцінка, зроблена для аналогічної фази циклу. Потім величина  $\widehat{a}_{1,t}$ , отримана за першим рівнянням, використовується для визначення нової оцінки коефіцієнта сезонності по другому рівнянню.

Прогноз наступного значення роду:

$$\widehat{x}_\tau(t) = \widehat{a}_{1,t} \widehat{f}_{t-l+\tau}, \quad \tau \leq l^t.$$

Величини  $\widehat{a}_{1,t}$  та  $\widehat{f}_t$  можуть бути записані через попередні умови та дані:

$$\begin{aligned} \widehat{a}_{1,t} &= \alpha_1 \sum_{n=0}^t (1 - \alpha_1)^n \frac{x_{t-n}}{\widehat{f}_{t-l-n}} + (1 - \alpha_1)^{t+1} \widehat{a}_{1,0}; \\ \widehat{f}_t &= \alpha_2 \sum_{n=0}^J (1 - \alpha_2)^n \frac{x_{t-nl}}{\widehat{a}_{1,t-nl}} + (1 - \alpha_2)^{J+1} \widehat{f}_{l,0}, \end{aligned}$$

де  $\widehat{a}_{1,0}$  – початкове значення  $a_1$ ;

$\widehat{f}_{l,0}$  – початкове значення  $f$  у відповідній  $i$  фазі циклу;

$J$  – найбільша ціла частина з  $\frac{t}{l}$ .

Отже, прогноз є функцією всіх минулих значень фактичного ряду, параметрів  $\alpha_1$  і  $\alpha_2$  на початкових умовах  $\widehat{a}_{1,0}, \widehat{f}_{1,0}, \widehat{f}_{2,0}, \dots, \widehat{f}_{l,0}$ .

*Модель сезонних явищ з лінійним зростанням*

Повна сезонна модель Уінтерса з лінійним зростанням аналогічна тільки що розглянутої:

$$\begin{aligned} \widehat{a}_{1,t} &= \alpha_1 \frac{x_t}{\widehat{f}_{t-l}} + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1}); \\ \widehat{f}_t &= \alpha_2 \frac{x_t}{\widehat{a}_{1,t}} + (1 - \alpha_2)\widehat{f}_{t-l}; \\ \widehat{a}_{2,t} &= \alpha_3(\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_3)\widehat{a}_{2,t-1} \quad 0 < \alpha_1, \alpha_2, \alpha_3 < 1; \end{aligned}$$

$$\widehat{x}_\tau(t) = (\widehat{a}_{1,t} + \tau \widehat{a}_{2,t}) \widehat{f}_{t-l+\tau}.$$

Єдиною змінною в вираженні для  $\widehat{a}_{1,t}$  є додавання  $\widehat{a}_{2,t-1}$  - найбільш пізньої оцінки адитивного фактора росту, що характеризує зміну середнього за повний сезонний цикл рівня процесу за одиницю часу (місяць). Вираз для поновлення коефіцієнта сезонності залишається тим же, що і раніше. Оцінки  $\widehat{a}_{2,t}$  модернізується за аналогічною процедурою експоненціального згладжування. Прогноз є тут функцією минулих і поточних даних, параметрів  $\alpha_1, \alpha_2, \alpha_3$  і початкових значень  $\widehat{a}_{1,0}, \widehat{a}_{2,0}, \widehat{f}_{l,0}$ . Якість і точність прогнозів залежить від цих чинників.

Оптимальні параметри  $\alpha_1, \alpha_2, \alpha_3$  Уінтерс пропонує знаходити експериментальним шляхом. Критерієм порівняння він бере стандартне відхилення помилки. При цьому передбачається, що прогноз не усунутий. Пошук, здійснювався за допомогою сітки значень  $\alpha_1, \alpha_2, \alpha_3$ . Функція стандартної помилки поблизу мінімуму передбачалася досить плоскою.

#### *Адитивна модель сезонних явищ*

Незважаючи на те що для економічних часових рядів мультиплікативна модель зазвичай виявляється найбільш підходящою, іноді потрібна адитивна модель. Розглянемо адитивну модель сезонних явищ з лінійним зростанням, запропоновану Г. Тейл і С. Вейдже.

Побудова такої моделі має на меті спрощення процедури прогнозування, оскільки комбінація мультиплікативної сезонної моделі з лінійним зростанням математично громіздка. Крім того, на практиці частіше зустрічаються експоненціальні тенденції, ніж лінійні. Тому заміна значень початкового часового ряду їх логарифмами перетворює експонентну тенденцію в лінійну і одночасно мультиплікативну сезонну модель в адитивну. Тоді тимчасової ряд можна представити таким чином:

$$x_t = a_{1,t} + g_t + \varepsilon_t ;$$

$$a_{1,t} = a_{1,t-1} + a_{2,t} ,$$

де  $a_{1,t}$  — величина рівня процесу після елімінування сезонних коливань;

$a_{2,t}$  — адитивний коефіцієнт росту;

$g_t$  — адитивний коефіцієнт сезонності;

$\varepsilon_t$  — білий шум.

Спочатку розглянемо адаптивну процедуру оновлення значень  $\widehat{a}_{1,t}$ . В момент  $t$  ми використовуємо спостереження  $x_t$ , про яке відомо, що

$$x_t = a_{1,t} + g_t + \varepsilon_t.$$

Однак про шум і сезонний чинник  $g_t$  ніякої інформації немає. Величину  $\varepsilon_t$  замінимо нулем, а в якості заміника для  $g_t$  візьмемо найостаннішу оцінку сезонного фактора  $g_{t-l}$ , де  $l$  - період сезонного циклу.

Величину  $x_t - \widehat{g}_{t-l}$  будемо розглядати як нове «Фактичне» значення  $a_{1,t}$ . Останньою оцінкою рівня  $a_1 \in \widehat{a}_{1,t-1}$ , але вона відповідає моменту  $t - 1$ , а не  $t$ . Тому необхідно до  $\widehat{a}_{1,t-1}$  додати ще  $\widehat{a}_{2,t}$ , але так як оцінку  $\widehat{a}_{2,t}$  ми ще не можемо отримати, то замість неї беремо оцінку  $\widehat{a}_{2,t-1}$  отриману на попередньому кроці. Це призводить до наступної процедури адаптації:

$$\widehat{a}_{1,t} = \alpha_1(x_t - \widehat{g}_{t-l}) + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1}),$$

яка при даних вагах  $\alpha_1$  та  $(1 - \alpha_1)$  оцінює  $a_{1,t}$  через найбільш свіже спостереження  $x_t$  і заздалегідь підрахованих величин  $\widehat{a}_{1,t-1}$ ,  $\widehat{a}_{2,t-1}$ ,  $\widehat{g}_{t-l}$ .

З огляду на рівняння  $a_{1,t}$  і віднімаючи з отриманого  $\widehat{a}_{1,t}$  колишню оцінку  $\widehat{a}_{1,t-1}$ , можемо отримати оцінку  $\widehat{a}_{2,t}$ . Однак оскільки обчислення  $\widehat{a}_{1,t}$  не є досконалими, зокрема тому, що не бралися до уваги залишки з рівняння  $x_t$ , то, очевидно, краще не покладатися на цю різницю повністю, а вважати її «фактичним» свідченням динаміки ряду і об'єднати зі старим значенням  $\widehat{a}_{2,t-1}$  за відомою формулою експоненціального згладжування

$$\widehat{a}_{2,t} = \alpha_2(\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_2)\widehat{a}_{2,t-1},$$

де  $\alpha_2$  та  $(1 - \alpha_2)$  — ваги двох середовищ інформації.

Нарешті ця ж процедура застосовується для отримання оцінки  $g_t$ . Нове «фактичне» значення сезонного фактора буде  $x_t - \widehat{a}_{1,t}$ , старе значення рівне  $\widehat{g}_{t-l}$ , експоненційно-згладженого значення

$$\widehat{g}_t = \alpha_3(x_t - \widehat{a}_{1,t}) + (1 - \alpha_3)\widehat{g}_{t-l}.$$

Всі три параметра згладжування будуть задовольняти умови  $0 < \alpha_1, \alpha_2, \alpha_3 < 1$ . Адаптивне прогнозування тепер провести порівняно просто. Припустимо, що  $t$  - поточний момент часу, так що  $\widehat{a}_{1,t}$ ,  $\widehat{a}_{2,t}$ ,  $\widehat{g}_t$ ,  $\widehat{g}_{t-1}$  є в нашому розпорядженні. Припустимо також, що ми хочемо отримати прогноз величини  $x_{t+\tau}$  (прогноз на  $\tau$  кроків вперед). Екстраполюємо тенденцію лінійного росту, використовуючи саме останнє значення коефіцієнта  $\widehat{a}_{2,t}$ , додаємо саму свіжу оцінку сезонного члена для цієї фази циклу і нехтуємо шумом. В результаті отримуємо

$$\widehat{x}_\tau(t) = \widehat{a}_{1,t} + \tau\widehat{a}_{2,t} + \widehat{g}_{t-l+\tau}$$

за умови, що  $0 < \tau \leq l$ . Якщо  $l < \tau \leq 2l$ , то необхідно  $\widehat{g}_{t-l+\tau}$  замінити на  $\widehat{g}_{t-2l+\tau}$ . Модель готова.

### *Альтернативи моделей*

Взагалі можливо безліч комбінацій різних типів тенденцій і циклічних явищ адитивного і мультиплікативного виду. В роботі представлені дев'ять можливих моделей, які узагальнено виражені однією формулою. Оскільки завжди необхідно використовувати модель, найбільш точно відображає динаміку процесу, то доцільно зупинитися на цьому питанні докладніше.

Дев'ять згаданих моделей складають три групи по три варіанти в кожній. Першої групу утворюють модель без тренда 1 - А, модель з адитивним лінійним трендом 1 - В і модель з мультиплікативним трендом 1-С. Другу групу складають три моделі з першої групи з накладеним на них адитивним сезонним ефектом. У третю групу входять три моделі з першої групи, але з накладеним на них мультиплікативним сезонним ефектом.

Всі дев'ять моделей можуть бути відображені в одній загальній записи

$$\widehat{a}_{1,t} = \alpha_1 d_1 + (1 - \alpha_1) d_2,$$

де  $\widehat{a}_{1,t}$  - поточний рівень ряду після елімінування сезонних коливань;

$\alpha_1$  - параметр згладжування,  $0 < \alpha_1 < 1$ ;

значення  $d_1$  і  $d_2$  дані в таблиці «Значення символів узагальнюючої формули», кожна клітина якої характеризує ту чи іншу модель.

### 2.3. Апроксимація полінеальних трендів за допомогою багатократних згладжувань

*Багаторазове згладжування* Поняття експоненційної середньої  $S_t$  для часового ряду  $x_t$  можна узагальнити на випадок експоненційних середніх більш високих порядків. Експоненціальна середня довільного  $p$ -го порядку визначається як

$$S_t^{[p]} = \alpha S_t^{[p-1]} + \beta S_{t-1}^{[p]},$$

де  $\beta = 1 - \alpha$ ;  $p = 1, 2, \dots, n$ ;  $S_t^{[0]} = x_t$ ;  $S_0^{[1]}, S_0^{[2]}, \dots, S_0^{[n]}$  - початкові значення експоненційних середніх відповідного порядку, тобто вирівнювання  $p$ -го порядку є простим експоненціальним згладжуванням, застосованим до результатів згладжування  $(p - 1)$ -го порядку.

Якщо в якості гіпотези тренда деякого процесу приймається поліном ступеня  $n$ , то метод експоненціального згладжування і прогнозування дозволяє обчислити коефіцієнти пророкує полінома через експоненціальні середні відповідних порядків.

У загальному випадку приймається гіпотеза, що досліджуваний процес є параболою  $n$ -го порядку, а прогноз на  $\tau$  кроків вперед виражається формулою

$$\widehat{x}_\tau(t) = \widehat{a}_1 + \widehat{a}_2 \tau + \frac{1}{2!} \widehat{a}_3 \tau^2 + \dots + \frac{1}{n!} \widehat{a}_{n+1} \tau^n,$$

де параметри  $\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_{n+1}$  потрібно визначити.

Ідея експоненціального передбачення заснована на тому, що прогноз здійснюється поліномом, які представляють перші  $n + 1$  членів розкладання процесу  $x_t$  в ряд Тейлора:

$$x_{t+\tau} = \sum_{k=0}^{\infty} \frac{\tau^k x_t^{(k)}}{k!} = \sum_{k=0}^n \frac{\tau^k x_t^{(k)}}{k!} = \sum_{k=0}^n \frac{\tau^k a_{k+1}}{k!}.$$

Якщо спостережуваний процес може бути представлений поліномом ступеня  $n$ , то всі похідні порядку  $n + 1$  і вище будуть дорівнюють нулю.

Фундаментальна теорема методу експоненціального згладжування і прогнозування, вперше доведена Р. Брауном і Р. Майером, говорить про те, що коефіцієнти  $\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_{n+1}$  пророкує полінома пов'язані з експонентними середніми

$$S_t^{[p]}, \quad p = 1, 2, \dots, n + 1$$

співвідношенням

$$S_t^{[p]} = \sum_{k=0}^n (-1)^k \frac{x_t^{(k)}}{k!} \frac{\alpha^p}{(p-1)!} \sum_{j=0}^{\infty} j^k \beta^j \frac{(p-1+j)!}{j!};$$

$$\beta = 1 - \alpha.$$

Таким чином, є  $n + 1$  рівняння, що дає згладжені значення  $S_t^{[p]}$  через лінійні комбінації похідних  $x_t^{[k]}$ .

Лінійні рівняння для згладжених величин можна записати в компактній формі, використовуючи матриці.

$$\text{Нехай } S_t \text{ буде вектором: } S_t = \begin{bmatrix} S_t \\ S_t^{[2]} \\ \vdots \\ S_t^{[p]} \end{bmatrix} \text{ і нехай } b_t \text{ буде вектором: } b_t = \begin{bmatrix} b_{1,t} \\ b_{2,t} \\ \vdots \\ b_{n+1,t} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{1!} x_t^{(0)} \\ \frac{1}{2!} x_t^{(1)} \\ \vdots \\ \frac{1}{n!} x_t^{(n)} \end{bmatrix}, \text{ коефіцієнтів в розкладі Тейлора.}$$

Тоді зміст фундаментальної теореми може бути виражено так:  $S_t = Mb_t$ ,

де  $M$  – матриця  $(n + 1) \cdot (n + 1)$  з елементами, що містять нескінченні суми ступенів постійної згладжування:

$$m_{ik} = \frac{\alpha^t}{(t-1)!} \sum_{j=0}^{\infty} j^k \beta^j \frac{(p-1+j)!}{j!}.$$

*Як стежити контрольний сигнал*

Для якнайшвидшого виявлення неадекватності моделі реального процесу, що необхідно для внесення відповідних змін до модель прогнозування, Р. Браун розробив метод аналізу прогнозуючої системи, що складається в підрахунку величини стежить контрольного сигналу. Стежить контрольний сигнал  $K_t$  визначається як сума помилок прогнозування  $e_t$  поділена на величину їх згладженого абсолютного значення:

$$\tilde{e}_t = (1 - \gamma)e_{t-1} + \gamma|e_t|,$$

тобто: 
$$K_t = \frac{\sum_{i=0}^t e_i}{(1 - \gamma)e_{t-1} + \gamma|e_t|},$$

де  $0 < \gamma < 1$  – стала згладжування.

Браун вказує значення критичних рівнів контрольного сигналу, перевищення яких говорить про необхідність додаткового вивчення і зміни моделі.

Такий прийом має два недоліки. По-перше, в разі коли контрольний сигнал вийшов за встановлені межі, він не обов'язково повернеться в ці ж межі, навіть якщо даний процес знову буде розвиватися за колишніми законами і прогнозуюче модель виявиться адекватною реальному процесу. Отже, потрібно втручання, щоб зробити суму помилок знову близькою до нуля і уникнути таким чином хибних сигналів тривоги. По-друге, можлива і зворотна ситуація, коли контрольний сигнал виходить із зазначених меж, а система починає давати більш точні прогнози.

## 2.4. Загальна модель Брауна

Р. Браун ще більш розширив можливості прогнозування з використанням адаптивних моделей. Він розглянув процес:

$$x_t = \xi_t + \varepsilon_t ,$$

де

$$\xi_t = a_1 f_1(t) + a_2 f_2(t) + \dots + a_n f_n(t) = \sum_{i=1}^n a_i f_i(t);$$

$\varepsilon_t$  – випадкові незалежні відхилення з середнім значенням 0 і дисперсією  $\sigma_t^2$ ;

$f_i(t)$  - деякі відомі, вибрані заздалегідь детерміновані функції від часу;

$a_i$  - коефіцієнти, що підлягають визначенню і адаптації.

Розглянутий процес може мати слабкі випадкові зміни одного або більше коефіцієнта. Р. Браун поставив собі за мету розробити адаптивну процедуру для перегляду оцінок коефіцієнтів при кожному отриманні нової фактичної точки ряду. Йому вдалося побудувати компактну ітеративну модель для випадку, коли функції, що входять в модель, задовольняють співвідношенню:

$$f(t) = Lf(t - 1),$$

де  $f(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix}$  – Вектор стовпчик із підібраних функцій;

$L$  – матриця постійних коефіцієнтів розмірність  $(n \cdot n)$  свого роду матриця переходу.

Такі функції є розв'язками лінійних різницевих рівнянь. Ними можуть бути лише поліноми, експоненти і синусоїди або їхні твори. Маючи матрицю переходу для відповідного набору функцій, використаних в моделі, необхідно також визначити значення функцій в будь-якої початковий момент часу, зазвичай при  $t = 0$  або  $t = 1$ . За вектору початкових значень і матриці  $L$  можна отримати значення  $f(t)$  для будь-якого моменту часу

$$f(t) = L^t f(0).$$

Для спрощення розрахунків Р. Браун за початок відліку часу приймає поточний момент  $T$ , тобто момент складання прогнозу, За критерій помилки, який слід мінімізувати, взята зважена сума квадратів відхилень:

$$\sum_{j=0}^T \beta^j \left[ x_{T-1} - \widehat{a}_T f(1-j) \right]^2,$$

$$\widehat{a}_T = [\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_n],$$

$\beta$  – дисконтіруючий фактор  $0 < \beta < 1$ .

Далі для функцій які ми підбираємо вводиться матриця  $(n \cdot n)$   $F(t)$ :

$$F_{ik} = \sum_{j=0}^t \beta^j f_i(-j) f_k(-j).$$

Для визначення оцінок коефіцієнтів моделі відповідно до обраного критерію Р.

Браун вивів  $n$ -компонентний вектор  $h = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} = F^{-1} f(0)$ , де  $F$  – матриця функцій в стабільному стані.

З його допомогою адаптивні коефіцієнти  $a_t$  повинні оновлюватися по формулі:

$$\widehat{a}_t = L' \widehat{a}_{t-1} + h e_1(t-1),$$

де  $L'$  – транспонована матриця  $L$ ;

$e_1(t-1)$  – помилка прогнозу розрахований в момент  $(t-1)$  на 1 крок вперед.

Константи  $h_t$  залежать тільки від приватного набору підбираються функцій, від вектора початкових значень  $f(0)$  і від величини  $\beta$ .

Очевидно, що при відсутності помилок прогнозування коефіцієнти будуть змінюватись по закону:

$$\widehat{a}_t = L' \widehat{a}_{t-1}.$$

Це пов'язано з переносом, початку відліку часу на кожному кроці на 1 інтервал вперед. Другий доданок в рівнянні  $\widehat{a}_t$  дає правило коригування коефіцієнтів в залежності від помилки прогнозу.

Досліджуючи дисперсію прогнозів, Р. Браун встановив, що моделі, що складаються з тригонометричних функцій і постійної складової, характеризуються приблизно однією і тією ж дисперсією прогнозів для всіх періодів прогнозування  $\tau$ . При грубої прикидки для цих моделей можна використовувати співвідношення:

$$D(\hat{x}_\tau) \approx \frac{1 - \beta^n}{2} \sigma_\varepsilon^2.$$

Отже, на практиці метод Р. Брауна завдяки своїй простоті може бути корисний на етапі грубих прикидок або для обробки великої кількості подібних рядів, коли проведення повної процедури ідентифікації та оцінки параметрів моделі.

### РОЗДІЛ 3. Адаптивна модель прогнозування часового ряду, генерованого авторегресивною схемою з дрейфуючими коефіцієнтами

У цьому розділі розглядається ще один метод прогнозування членів ряду на основі попередніх даних, що називається адаптивною фільтрацією. Цей Метод застосовується, коли поточний член ряду пов'язаний з попередніми членами цього ж ряду лінійним співвідношенням, якому коефіцієнти можуть бути змінними.

#### 3.1. Загальна схема адаптивного фільтру

У роботах С. Уілрайт і С. Макрідакіс спробували використовувати адаптивну фільтрацію для отримання короткострокових прогнозів. Однак Д. Монтгомері піддав їх метод критиці; метод охарактеризований як призначений для обробки лише авторегресійних стаціонарних рядів. Параметри авторегресії у методі оцінюються в ітеративному процесі навчання моделі. Значення параметра, що забезпечує збіжність цього процесу, знаходиться шляхом проб. Таке оцінювання з обчислювальної точки зору значно поступається прямому оцінюванню методом лінійної множинної регресії. Крім того, на тому ж ряду місячних продажів шампанського у Франції, що використовувався при розрахунках авторами методу, було проведено прогнозування моделі П. Уінтерса і моделі Дж. Бокса - Г. Дженкінса. Перевага залишилася за двома останніми моделями.

Враховуючи досвід перших спроб побудови моделей, заснованих на ідеї адаптивної фільтрації, нам пропонують спосіб організації адаптації моделей авторегресійного типу на принципах, подібних до тих, які використовуються в експоненційній середній.

Операція лінійної фільтрації полягає у обчисленні виваженої суми попередніх спостережень. У найбільш загальному вигляді ця сума записується як

$$S = \sum_{i=t-l+1}^t \omega_i x_i$$

де  $S$  – виважена середня;

$\omega_i$  – вага, приписана спостереженню  $i$ ;

$x_i$  – значення, яке спостерігаємо момент  $i$ ;

$l$  – число спостережень, що використовується при підрахунку  $S$ ;

Метод рухомих середніх, наприклад, полягає в тому, що середня  $l$  останніх членів ряду, а потім отримане середнє використовується як прогноз, тобто:

$$\widehat{x}_1(t) = S = \frac{1}{l} (x_t + x_{t-1} + \dots + x_{t-(l-1)}).$$

Таким чином, останні  $l$  даних мають однакову вагу  $\frac{1}{l}$ , а більш рані дані мають нульову вагу.

Формула експонентного згладжування:

$$S_{t+1} = \alpha x_t + (1 - \alpha)S_t,$$

Може бути записаним у вигляді:

$$S_{t+l} = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + \dots$$

Тут найбільша вага має поточне значення  $x_t$ , а для більш старих спостережень має місце геометричне зменшення ваги.

Звідси видно, що ці методи відрізняються лише правилом визначення ваги  $\omega_i$ . Метод прогнозування з допомогою адаптивного фільтра це інший підхід до визначення ваги, інший спосіб побудови фільтра. Очевидно, що через відмінності у терезах від цих методів слід очікувати різних результатів, різної точності прогнозів.

Перша робота, присвячена синтезу фільтра, була опублікована М. Вінером у 40-ві роки. Вінер розглянув питання побудови лінійних фільтрів з постійними коефіцієнтами для елімінування шуму та для згладжування та прогнозування стаціонарних процесів. Розроблені ним процедури дають оптимальні з погляду критерію найменших квадратів результати для випадків, коли ряд справді стаціонарний.

Продовжуючи роботу Вінера, різні автори, включаючи Р. Кальмана і Р. Бьюсі, розробили процедури, що дають оптимальні лінійні фільтри, що змінюються в часі, для нестационарних часових рядів. Для таких рядів метод Кальмана – Бьюсі може дати кращі результати, ніж звичайний метод Вінера.

Недолік процедур Вінера і Кальмана - Бьюсі, в тому, що фільтри повинні бути побудовані на основі апріорних даних або припущень щодо статистики часового ряду. На практиці ці два методи фільтрації дають мінімальні помилки прогнозу. За умови, що статистичні характеристики ряду справді відповідають апріорній інформації, на основі якої було побудовано фільтри.

У методі адаптивної фільтрації статистики ряду не вимірюються, але враховуються неявно в процесі побудови фільтра та ітеративного оновлення ваги під час його адаптації.

### 3.2. Адаптація коефіцієнтів моделей авторегресії

Отже, ключем до ефективності адаптивної фільтрації є правило, що використовується для адаптації ваги на кожному кроці. Це може бути виведено шляхом аналізу помилки прогнозу.

У моделі адаптивного фільтра передбачається, що прогноз наступного члена ряду може бути отриманий за допомогою виваженої суми  $l$  колишніх членів ряду:

$$\hat{x}_1(t) = S,$$

де  $S$  – виважена середня, визначення виразом з пункту (3.1.).

Цей вираз має авторегресивний характер. Тому цю модель також будемо назвати адаптивною моделлю авторегресії порядку  $l$ . Після того як стає відомим  $x_{t+1}$  – фактичне значення ряду в момент  $t + 1$ , можна підрахувати похибку прогнозу  $e_{t+1} = x_{t+1} - \hat{x}_1(t) = x_{t+1} - \sum \omega_i x_i$ .

Всі суми тут і надалі будемо рахувати від  $t - l + 1$  до  $t$ .

Запишемо вираз для квадрату похибки:

$$e_{t+1}^2 = \left( x_{t+1} - \sum \omega_i x_i \right)^2 = x_{t+1}^2 - 2 \sum_i \omega_i x_i x_{t+1} + \sum_i \sum_j \omega_i \omega_j x_i x_j.$$

Цей вираз показує, що квадрат похибки є квадратичною функцією від ваги  $\omega_i$ . Геометричною інтерпретацією цієї функції у тривимірному просторі є параболічний циліндр. При певному поєднанні ваг (вздовж прямої  $\sum \omega_i x_i = x_{t+1}$ ) функція набуває мінімального значення, що дорівнює нулю.

Передбачається, що в момент  $t - 1$  були отримані деякі оцінки вагових коефіцієнтів, за якими підраховано оцінку  $x_t$ . Якщо використані оцінки вагових коефіцієнтів призвели до ненульової помилки, тобто є можливість скоригувати ваги так, щоб зменшити помилку. Можна було б вибрати такі ваги, які б зводили помилку до нуля. Але треба мати на увазі, що за оцінками ваг, отриманими в момент  $t$ , буде сформована оцінка величини  $x_{t+1}$ . Тим часом у момент  $t + 1$  функція  $e_{t+1}^2$  від ваги зміниться, оскільки її коефіцієнти (тобто значення  $x - i$ ) будуть іншими. У тривимірному просторі цю функцію відобразить інший параболічний циліндр, що приймає мінімальне значення вздовж нової прямої. Таким чином, процедуру адаптації ваги в момент  $t$  не можна розглядати ізольовано, так як коригування ваг на основі поточної помилки проводиться з метою мінімізації майбутньої помилки. Тому процедура адаптації ваги повинна враховувати динамічні особливості процесу і передбачати згладжування суто випадкових коливань параметрів системи, що вивчається.

Очевидно, коригування ваг можна здійснювати різними способами. Ми розглянемо один із них. Для адаптації ваг скористаємося методом якнайшвидшого спуску. Сутність його полягає у виборі початкової точки на досліджуваній поверхні та подальшому пересуванні до нижчої точки поверхні із застосуванням ітеративної процедури. Для цього необхідно мати можливість обчислювати в кожній точці поверхні вектор, що вказує напрямку руху. Тоді можна коригувати ваги та-

ким чином, що нові ваги будуть представляти точку, яка ближче до оптимального набору ваг у порівнянні зі старими вагами. При використанні методу найшвидшого спуску коригування ваг здійснюється за таким правилом:

$$W_H = W_C - k \operatorname{grad}(e_{t+1}^2),$$

де  $W_C$  – вектор старих ваг;

$W_H$  – вектор нових ваг;

$k$  – коефіцієнт ( $k > 0$ );

$\operatorname{grad}(e_{t+1}^2)$  – вектор, градієнт  $e_{t+1}^2$ .

Це рівняння показує, що коригування терезів здійснюється шляхом додавання до старого вектора вагів поправки, одержуваної множенням коефіцієнта  $k$  на градієнт, взятий з негативним знаком. Градієнт із негативним знаком вказує найкоротший шлях досягнення мінімуму досліджуваної поверхні, а коефіцієнт визначає, наскільки ми просуваємося в цьому напрямку.

Компоненти градієнта знаходимо диференціюванням квадрата похибки за вагами:

$$\frac{\partial e_{t+1}^2}{\partial \omega_i} = 2e_{t+1} \frac{\partial e_{t+1}}{\partial \omega_i} = -2e_{t+1}x_i.$$

В результаті отримуємо, що градієнт в цілому

$$\operatorname{grad}(e_{t+1}^2) = -2e_{t+1}X,$$

де  $X$  – вектор  $l$  останніх спостережень. Це визначає спосіб коригування ваг:

$$W_H = W_C + 2ke_{t+1}X.$$

Невідомим залишається значення  $k$ , що визначає швидкість руху у напрямку, зворотному до градієнта. Щоб визначити характер впливу  $k$  на процедуру адаптації, отримаємо таким чином. Повернемося в точку  $t$  і знову зробимо прогноз, але вже з вагами  $W_H$ . Отримаємо нове значення похибки:

$$(e_H)_{t+1} = x_{t+1} - \sum[(\omega_c)_i + 2ke_{t+1}x_i]x_i = x_{t+1} - \sum(\omega_c)_i x_i - \sum 2ke_{t+1}x_i^2 = e_{t+1} - 2ke_{t+1} \sum x_i^2 = e_{t+1} (1 - 2k \sum x_i^2),$$

де  $e_{t+1}$  – помилка, отримана при старих вагах  $(\omega_c)_i$ .

Тепер якщо позначимо, що:

$$k = \frac{\alpha}{2 \sum x_i^2}, \quad \text{то } (e_H)_{t+1} = e_{t+1}(1 - \alpha) \text{ і при } 0 < \alpha < 2 \quad |(e_H)_{t+1}| < |e_{t+1}|.$$

Отже,  $\alpha$  визначає реакцію моделі на отриману помилку і коригує ваги так, щоб компенсувати помилку на  $(1 - |1 - \alpha|) \cdot 100\%$  або за  $0 < \alpha < 1$  на  $\alpha \cdot 100\%$ . Згадаймо просте експоненційне згладжування, де експоненційна середня коригується аналогічним чином:  $S_t = S_{t-1} + \alpha e_t$ , де  $e_t = x_t - S_{t-1}$ .

Будемо називати  $\alpha$  параметром адаптації моделі і вважати його заданим і постійним,  $k$  при цьому буде змінною величиною. Оптимальне значення  $\alpha$ , що забезпечує мінімум середнього квадрата помилки, можна визначити методом проб на ретроспективному матеріалі. Визначення оптимального значення можна назвати процедурою «навчання» адаптивного фільтра.

## РОЗДІЛ 4. Доведення коректності алгоритму на прикладі

### 4.1. Використання алгоритму в EXEL

#### *Постановка задачі*

2021 рік виявився невдалим для столичних забудовників. Сподіваючись наростити продажі після кризового 2020-го, девелоперським компаніям довелося спостерігати за відпливом покупців на фоні рекордного зростання цін.

За рік вартість квадратного метра київської новобудови зросла на 20-30% до майже 1 300 дол. Востаннє на такому високому рівні ціни були у 2013 році.

Протягом року вартість житла підіймалася нерівномірно через різні фактори: від особливостей діяльності держорганів до роботи "друкарського верстата" США.

Побудуємо прогноз по моделі Брауна щоб спрогнозувати подальшу поведінку ринку нерухомості. Відомо, що ціни протягом 2021 року були такими, дані наведені в Таблиці 4.1.1:

Таблиця 4.1.1

2021 грн/м <sup>2</sup>	січень	лютий	березень	квітень	травень	червень	липень	серпень	вересень	жовтень	листопад	грудень
фактичні угоди	35720	36272	36974	37439	37173	37531	38740	40432	40795	42104	42447	43214

Дані взято з сайту: [Ціни на квартири в Києві вирости за рік на третину\\_ який прогноз на 2022\\_ Економічна правда.html](#)

Вихідний часовий ряд містить 12 рівнів спостереження показника  $Y(t)$ .

#### *Методологія*

Модель Брауна може відображати розвиток не тільки у вигляді лінійної тенденції, але також у вигляді випадкового процесу, що не має тенденції, а також у вигляді мінливої тенденції. Відповідно розрізняють моделі Брауна:

- *Нульового порядку*, яка описує процеси, які не мають тенденцій розвитку.

Вона має один параметр  $A_0$  (оцінка поточного рівня). Прогноз розвитку на  $k$  кроків уперед здійснюється з формулою  $Y(t+k) = A_0$ ;

- *Першого порядку* ( $Y(t+k) = A_0 + A_1k$ ). Коефіцієнт  $A_0$  – значення, близьке до останнього рівня, і представляє як закономірну складову цього рівня. Коефіцієнт  $A_1$  визначає приріст, що сформувався в основному до кінця періоду спостережень, та відображає швидкість росту на більш ранніх етапах;
- *Другого порядку*, що відображає розвиток у вигляді парабольної тенденції з мінливою «швидкістю» і «прискоренням». Вона має три параметри ( $A_2$  – оцінка поточного приросту або «прискорення»). Прогноз реалізується за формулою:

$$Y(t+k) = A_0 + A_1k + A_2k^2.$$

Розглянемо етапи побудови лінійної адаптивної моделі Брауна.

*Етап 1.* По першим п'ятьом точкам часового ряду оцінюються початкові значення  $A_0$  та  $A_1$ , параметрів моделі за допомогою методу найменших квадратів для лінійної апроксимації:

$$Y_p(t) = A_0 + A_1t.$$

*Етап 2.* З використанням параметрів  $A_0$  та  $A_1$  по моделі Брауна знаходимо прогноз на один крок ( $k = 1$ ):

$$Y_p(t, k) = A_0(t) + A_1(t)k.$$

*Етап 3.* Розрахункове значення  $Y_p(t, k)$  економічного показника порівнюють з фактичним  $Y(t)$  і обраховується величина їх розбіжностей. При  $k = 1$  маємо:

$$e(t+1) = Y(t+1) - Y_p(t, 1).$$

*Етап 4.* Відповідно до цієї величини корелюються параметри моделі. У Моделі Брауна модифікація здійснюється в такий спосіб:

$$A_0(t) = A_0(t-1) + A_1(t-1) + (1-\beta)^2 e(t);$$

$$A_1(t) = A_1(t-1) + (1-\beta) e(t),$$

де  $\beta$  – коефіцієнт дисконтування даних, що змінюється в межах від 0 до 1, який характеризує знецінення даних за одиницю часу, а також відображає ступінь довіри більш пізнім спостереженням. Оптимальне значення  $\beta$  знаходиться ітеративним шляхом, тобто багаторазовою побудовою моделі при різних  $\beta$  і вибором найкращої, або по формулі:

$$\beta = \frac{N - 3}{N - 1},$$

де  $N$  – довжина часового ряду,  $e(t)$  – помилка прогнозування.

*Етап 5.* По моделі із скоректованими параметрами  $A_0$  та  $A_1$  знаходять прогноз на наступний момент часу. Повертаємося в пункт 3, якщо  $t < N$ . Якщо  $t = N$ , то побудовану модель можна використовувати для прогнозування на майбутнє.

*Етап 6.* Інтервальний прогноз будується як для лінійної моделі кривої росту.

#### *Результати дослідження*

Скористаємося схемою адаптивного прогнозування. Початкові оцінки параметрів одержимо по першим п'ятьом точкам за допомогою МНК по формулах:

$$A_1 = \frac{\sum(t - t_{cp})Y(t) - Y_{cp}}{\sum(t - t_{cp})^2}$$

де  $t_{cp}$  – середнє значення фактору «час»;

$Y_{cp}$  – середнє значення досліджуваного показника.

#### *Оцінка початкових значень параметрів моделі*

	t	Y(t)	(t-tcp)^2	Y(t)-Ycp	t-tcp	(t-tcp)(Y(t)-Ycp)	
tcp	1	35720	4	-995,6	-2	1991,2	
	3	36272	1	-443,6	-1	443,6	
	3	36974	0	258,4	0	0	
Ycp	4	37439	1	723,4	1	723,4	
	36715,6	5	37173	4	457,4	2	914,8

Далі я намагаюсь застосувати модель Брауна напряму до фактичних значень щоб підрахувати прогнозовані. За такими формулами:

$$A_t = \alpha Y(t) + (1 - \alpha)A_{t-1}$$

$$A_t'' = \alpha A_0 + (1 - \alpha)A_{t-1}''$$

$$a_t = 2A_t - A_t''$$

$$b_t = \left( \frac{\alpha}{1 - \alpha} \right) (A_t - A_t'')$$

$$F_{t+m} = a_t + b_t m$$

На першому кроці ініціалізації формули набувають вигляд:

$$A_1 = A_1'' = Y(t)_1, \quad a_1 = Y(t)_1$$

$$b_1 = \frac{(Y(t)_2 - Y(t)_1) + (Y(t)_4 - Y(t)_3)}{2}$$

Оцінка параметрів моделі Брауна

Таблиця 4.1.3

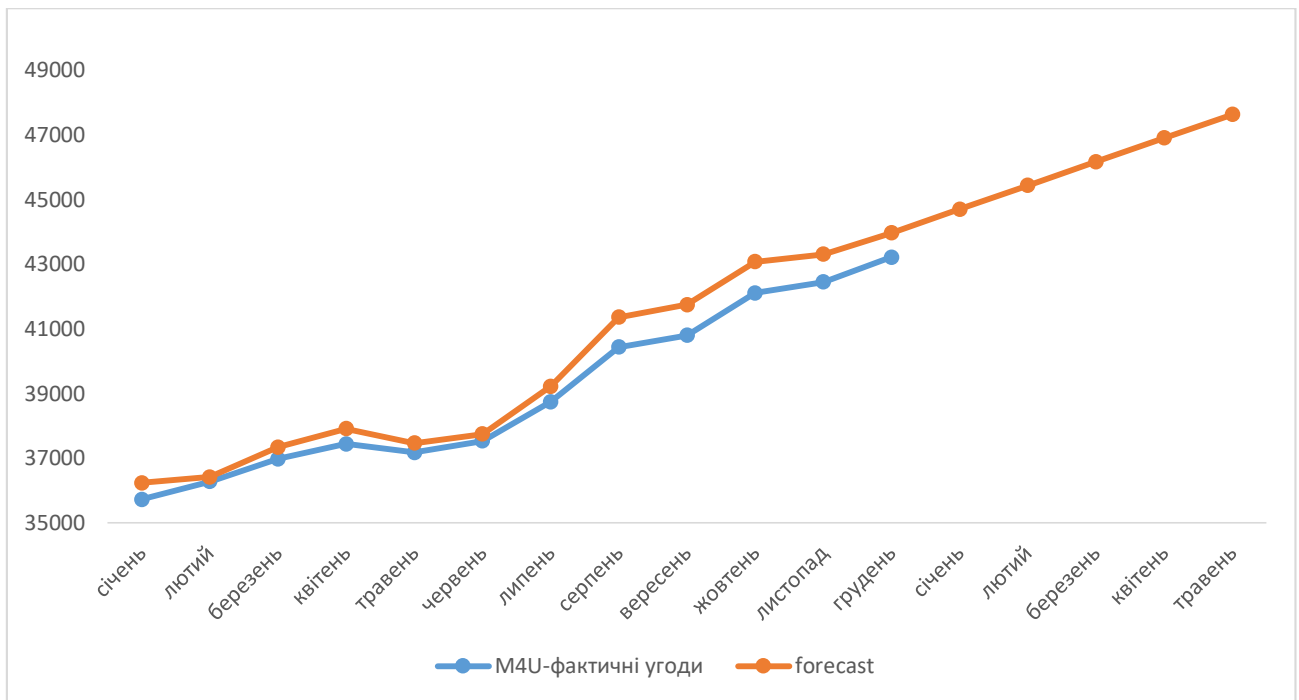
2021 грн/м <sup>3</sup> t	Y(t)	A <sub>0</sub>	A <sub>1</sub>	a <sub>t</sub>	b <sub>t</sub>	forecast	e(t)	e(t) <sup>2</sup>	
січень	1	35720	35720	35720	35720	508,5			
лютий	2	36272	36067,17	35938,34	36195,99	218,34	36228,5	43,5	1892,25
березень	3	36974	36637,49	36378,05	36896,93	439,7137	36414,33	559,6696	313230
квітень	4	37439	37141,58	36858,25	37424,91	480,1983	37336,65	102,352	10475,93
травень	5	37173	37161,34	37048,87	37273,81	190,6191	37905,1	-732,105	535977,2
червень	6	37531	37393,83	37265,82	37521,83	216,9509	37464,43	66,57116	4431,72
липень	7	38740	38240,47	37878,8	38602,13	612,9753	37738,78	1001,216	1002433
серпень	8	40432	39618,77	38973,11	40264,44	1094,308	39215,11	1216,891	1480823
вересень	9	40795	40358,53	39844,43	40872,63	871,323	41358,74	-563,744	317807,8
жовтень	10	42104	41456,29	40858,17	42054,42	1013,739	41743,95	360,0502	129636,1
листопад	11	42447	42079,37	41626,21	42532,53	768,0427	43068,16	-621,16	385840
грудень	12	43214	42792,96	42360,01	43225,92	733,7983	43300,58	-86,5755	7495,32

Наступний мій крок полягає в тому, що я хочу спрогнозувати цінову політику ринку на період січень-червень 2022, до цього для прогнозування я використовувала формулу  $F_t = a_t + b_t$ , надалі формула змінюється відповідно до врахування значення параметру  $m$  – наступного кроку, який дорівнює різниці наступного кроку від загальної кількості періодів, яких у нас становить 12, тобто:

$$F_{t+m} = a_t + b_t m$$

січень	13						43959,72		
лютий	14						44693,52		
березень	15						45427,32		
квітень	16						46161,11		
травень	17						46894,91		
червень	18						47628,71		

Надалі я хочу визначити значення параметру  $\alpha$  – який б був для мене самим оптимальним, для цього я рахую  $SSE = \sum_{i=1}^{12} e(t)^2$ . В Excel є функція – «Розв’язувач», нею я і скористалась щоб мінімізувати  $SSE$ , змінюючи параметр  $\alpha$ , додавши деякі обмеження, що:  $0,01 \leq \alpha \leq 0,96$ . Таким чином я отримала, що  $\alpha = 0,628922492309033$ . Задля кращого розуміння я побудувала графік (додаток А):



Графік 4.1.1 (співвідношення між фактичною ціною за 1 м<sup>2</sup>, до прогнозованої)

## 4.2. Інсталювання в програмному кодї середовища Python

Я вирішила продовжити дослідження моделі Брауна, цього разу я буду інсталювати метод в середовищі Python. Для цього були використані такі бібліотеки:

```
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.api import SimpleExpSmoothing
import csv
```

Бібліотека pandas – програмна бібліотека, написана для мови програмування Python для маніпулювання даними та їхнього аналізу. Вона, зокрема, пропонує структури даних та операції для маніпулювання чисельними таблицями та часовими рядами.

Бібліотека `matplotlib.pyplot` - бібліотека на мові програмування Python для візуалізації даних двовимірною 2D графікою (3D графіка також підтримується). Отримані зображення можуть бути використані як ілюстрації в публікаціях.

Бібліотека `statsmodels.tsa` - містить модельні класи та функції, які корисні для аналізу часових рядів. До базових моделей належать одновимірні авторегресивні моделі (AR), векторні авторегресійні моделі (VAR) та одновимірні моделі авторегресії ковзного середнього (ARMA). Нелінійні моделі включають динамічну регресію з перемиканням Маркова та авторегресію. Він також включає описову статистику для часових рядів, наприклад автокореляцію, функцію часткової автокореляції та періодограму, а також відповідні теоретичні властивості ARMA або пов'язаних процесів.

`SimpleExpSmoothing` - Просте експоненційне згладжування.

Модуль `csv` дає програмісту можливість виконувати структурний аналіз файлів CSV (Comma Separated Values – змінні, розділені комами). Файл CSV – це текстовий файл, у якому кожен рядок має кілька полів, розділених комами, або іншими роздільниками. Ви можете розглядати кожен рядок як ряд, а кожне поле як стовпець. Формат CSV не має стандарту, але ці файли досить схожі, так що модуль CSV може розпізнавати більшу частину цих файлів. Ви також можете створювати файли CSV, використовуючи цей модуль.

Далі я зчитую данні з мого csv файлу:

```
index_file, data = [], []
```

```
file_data = open('realtyKyiv.csv')
csvreader = csv.reader(file_data)
```

```
for row in csvreader:
    index_file.append(row[0])
    data.append(int(row[1]))
```

```
file_data.close()
```

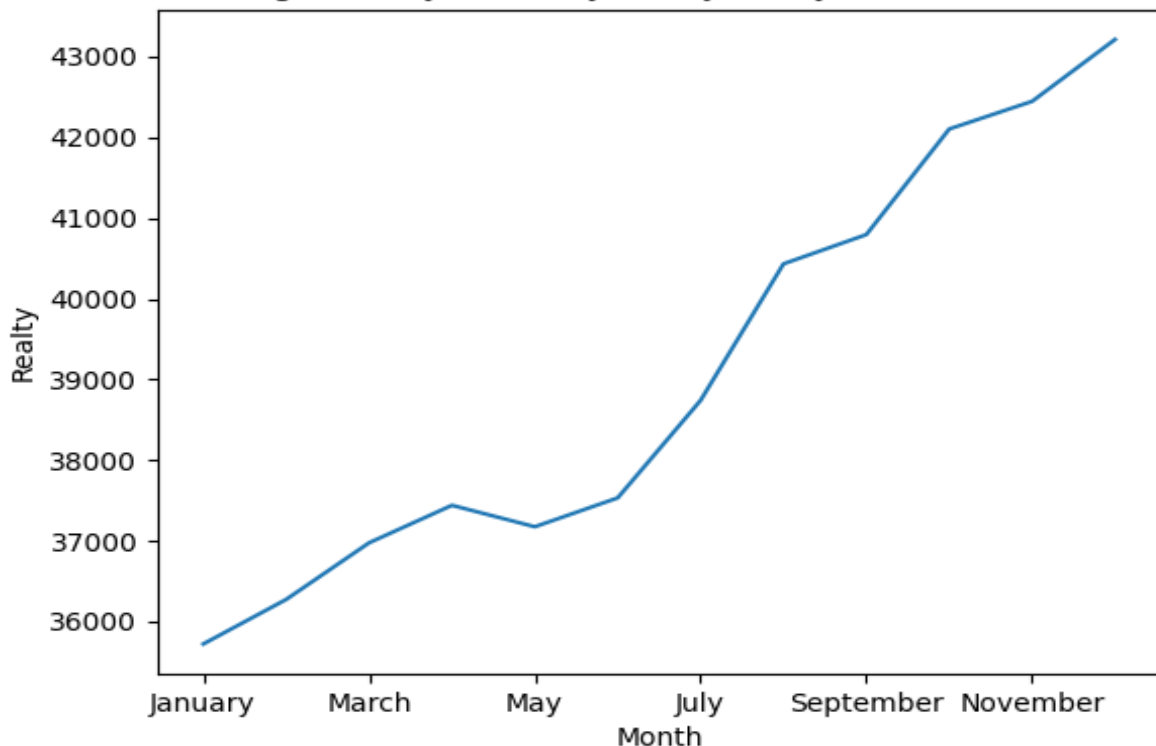
```
realty_data = pd.Series(data, index_file)
```

Файлі використовуються дані *Таблиці 4.1.1*, наведеної в пункті 4.1.

Зчитавши їх, я їм надаю значення змінної і будую графік (В кращій якості представлений в Додатку Б)

```
ax = realty_data.plot()
ax.set_xlabel("Month")
ax.set_ylabel("Realty")
plt.title("Figure 1: Kyiv's realty from January to December.")
plt.show()
index = pd.date_range(start='2021-01', freq='MS',
periods=12)
realty_data = pd.Series(data, index)
```

Figure 1: Kyiv's realty from January to December.



Графік 4.2.1(зміна фактичної ціни за 2021 рік)

Тепер ми можемо безпосередньо перейти до програмування моделі Брауна, або як його ще називають просте експоненційне згладжування.

Тут ми запускаємо три варіанти простого експоненційного згладжування:

1. `fit1` Ми не використовуємо автоматичну оптимізацію, а натомість вибираємо явно надати моделі з  $\alpha = 0.2$  параметр
2. `fit2` Як вище, вибираємо  $\alpha = 0.6$
3. `fit3` Ми дозволяємо `statsmodels` автоматично знаходити оптимізовані  $\alpha$  які становлять цінність для нас. Це рекомендований підхід.

Так як модель Брауна досліджує процеси з простою динамікою без тренду і сезонності, то алгоритм виглядатиме наступним чином:

```
fit1 = SimpleExpSmoothing(realty_data,
initialization_method="heuristic").fit(smoothing_level=0.
2, optimized=False)
fcast1 = fit1.forecast(3).rename(r"$\alpha=0.2$")
print(f"alpha=0.2: {fcast1.values}.")

fit2 = SimpleExpSmoothing(realty_data,
initialization_method="heuristic").fit(smoothing_level=0.
6, optimized=False)
fcast2 = fit2.forecast(3).rename(r"$\alpha=0.6$")
print(f"alpha=0.6: {fcast2.values}.")

fit3 = SimpleExpSmoothing(realty_data,
initialization_method="estimated").fit()
fcast3 = fit3.forecast(3).rename(r"$\alpha=%s$" %
fit3.model.params["smoothing_level"])
print(f"alpha={fit3.model.params['smoothing_level']}:
{fcast3.values}.")
```

Отримаємо таку ціну:

	$\alpha=0,2$	$\alpha=0,6$	$\alpha=0.9999999850988388$
січень	34610	34610	35720
лютий	34860	35280	35730
березень	35190	35900	36270
квітень	35490	36530	36970
травень	35900	37050	37450
червень	36220	37150	37200
липень	36470	37370	37550
серпень	36940	38190	38740

вересень	37610	39530	40430
жовтень	38200	40300	40790
листопад	38990	41380	42130
грудень	39700	42000	42480

Тоді ми виконуємо побудову графіку:

```
plt.figure(figsize=(12, 8))
plt.plot(realty_data, marker="o", color="black")

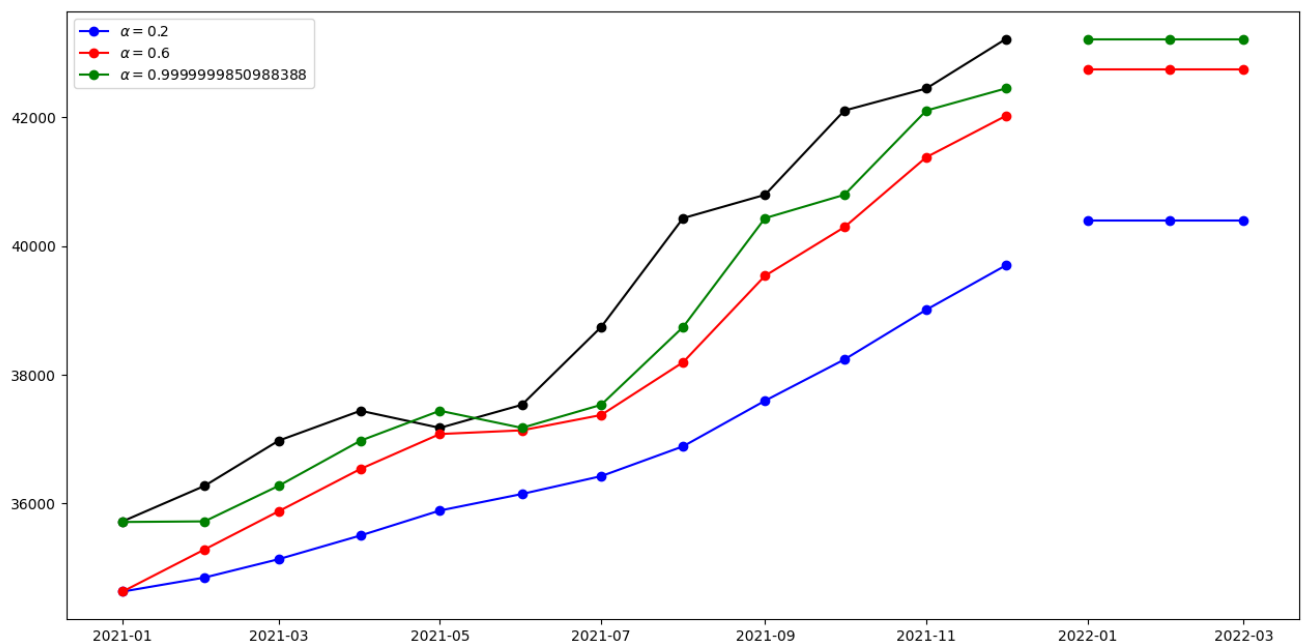
plt.plot(fit1.fittedvalues, marker="o", color="blue")
(line1,) = plt.plot(fcast1, marker="o", color="blue")

plt.plot(fit2.fittedvalues, marker="o", color="red")
(line2,) = plt.plot(fcast2, marker="o", color="red")

plt.plot(fit3.fittedvalues, marker="o", color="green")
(line3,) = plt.plot(fcast3, marker="o", color="green")
plt.legend([line1, line2, line3], [fcast1.name,
fcast2.name, fcast3.name])

plt.show()
```

Який матиме вигляд (детальніше можна розглянути в Додатку В):



Графік 4.2.2 (зміна прогнозованих цін відносно фіксованої, залежно від значення  $\alpha$ )

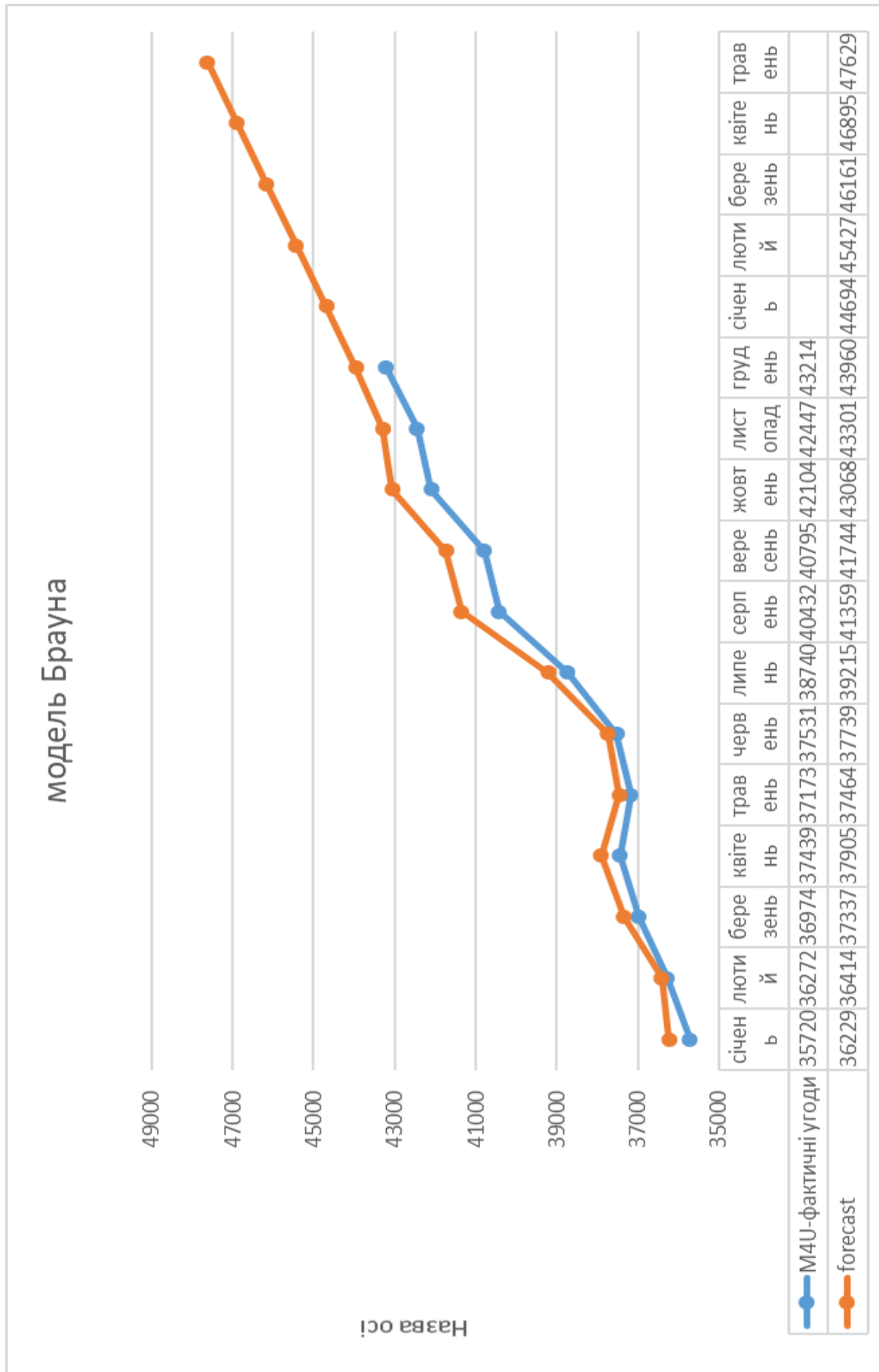
## Висновки

Закінчуючи розгляд адаптивних методів прогнозування, відзначу їх деякі особливості. Суть адаптації полягає в тому, що модель слідує за процесом. Це обумовлює відставання змін в моделі від нових тенденцій в реальному процесі, і чим більше час попередження, тим більше невідповідність між прогнозом і фактичним значенням ряду. Отже, моделі розглянутого класу можна рекомендувати для отримання в основному короткострокових прогнозів.

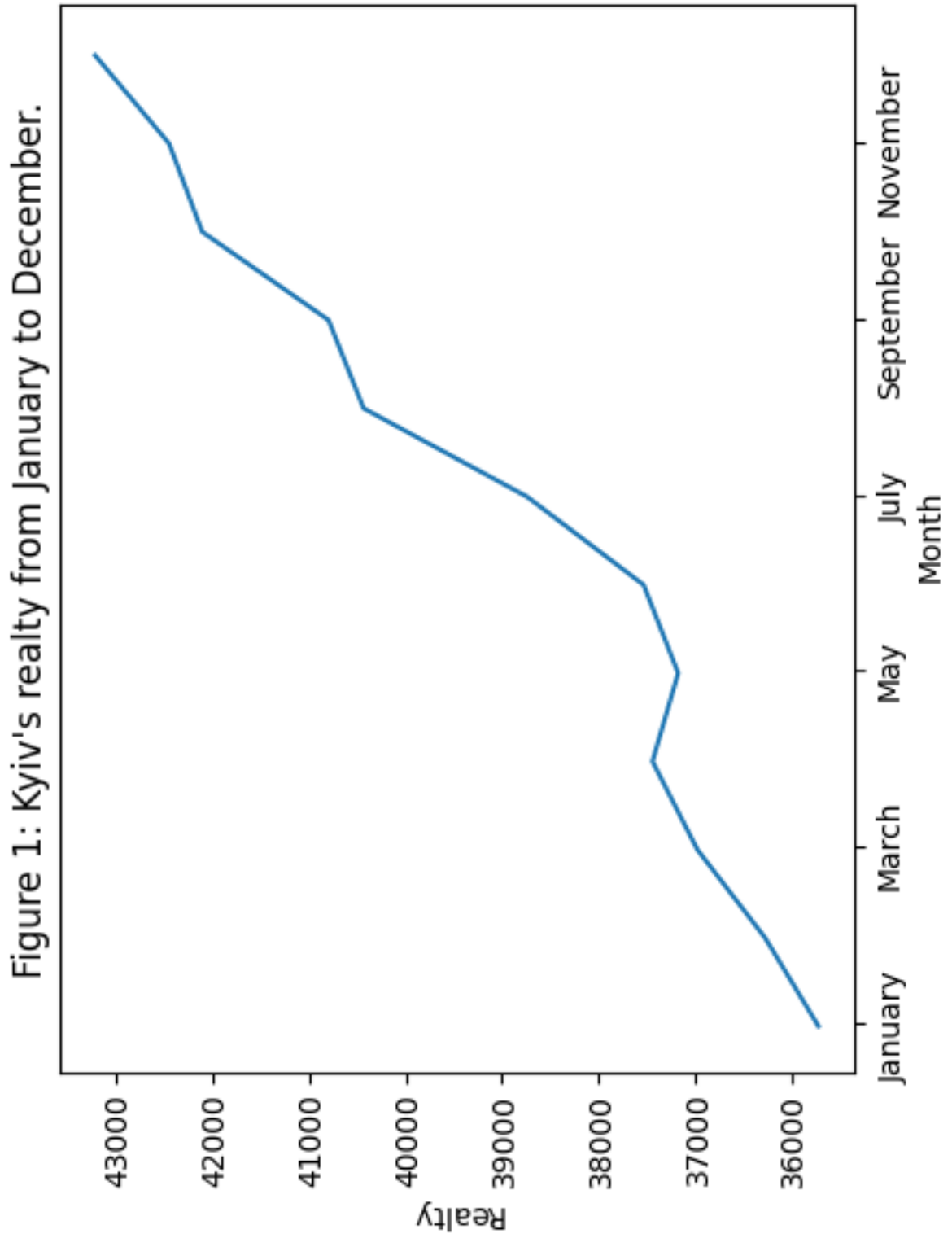
Багато з розглянутих моделей характеризують зв'язок між досліджуваною величиною і часом. Ця обставина сама по собі є досить серйозним обмеженням. З іншого боку, час в моделі висловлює еволюцію всього комплексу умов протікання процесу, через час вихідний ряд неявно пов'язаний з безліччю взаємопов'язаних факторів, врахувати вплив яких порізно важко. За рахунок спрощеного уявлення досліджуваної величини, пов'язаної з одним лише фактором часу, моделювання стає можливим навіть при самій мізерній інформації. Позитивною рисою адаптивних методів є те, що з їх допомогою ретельно вивчається внутрішня структура тимчасового ряду, взаємозв'язок його послідовних членів, а моделі, які є інструментом прогнозу, чуйно реагують на динамічні зміни і відповідно перебудовуються тим чи іншим чином, з огляду на знецінення застарілої інформації.

Сфера застосування адаптивних моделей одновимірного ряду досить широка. Вони можуть бути використані для прогнозування попиту і пропозиції, кон'юнктурних коливань фінансового ринку, окремих економічних і техніко-економічних показників, рівня запасів в системах матеріально-технічного постачання, для прогнозування структурних і технологічних зрушень, для визначення траєкторій деяких глобальних показників.

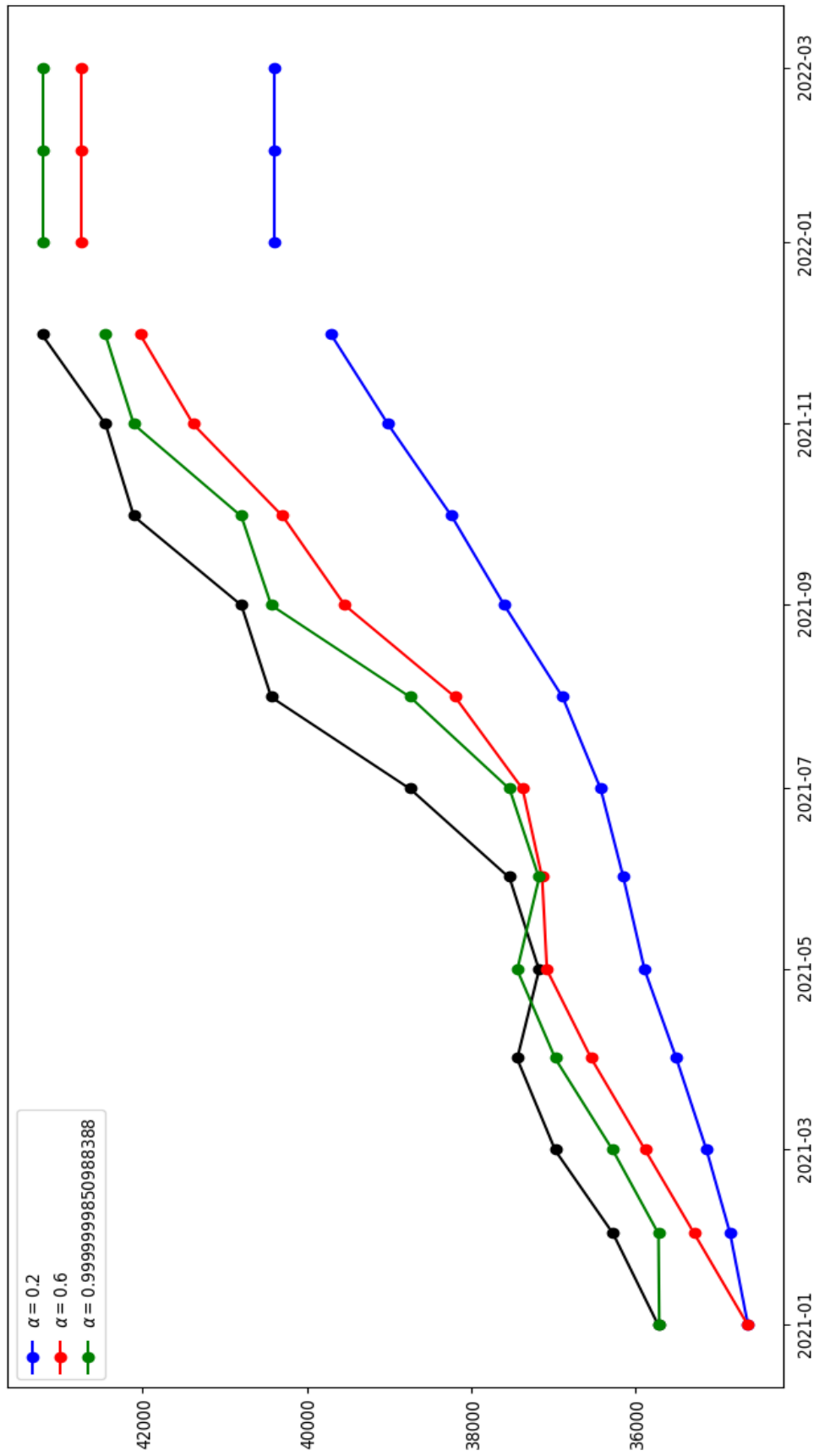
## Додаток А



## Додаток Б



## Додаток В



## Додаток Г

```

import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.api import SimpleExpSmoothing
import csv

index_file, data = [], []

file_data = open('realtyKyiv.csv')
csvreader = csv.reader(file_data)

for row in csvreader:
    index_file.append(row[0])
    data.append(int(row[1]))

file_data.close()

realty_data = pd.Series(data, index_file)

ax = realty_data.plot()
ax.set_xlabel("Month")
ax.set_ylabel("Realty")
plt.title("Figure 1: Kyiv's realty from January to
December.")
plt.show()

# Simple Exponential Smoothing
index = pd.date_range(start='2021-01', freq='MS',
periods=12)
realty_data = pd.Series(data, index)

fit1 = SimpleExpSmoothing(realty_data,
initialization_method="heuristic").fit(smoothing_level=0.
2, optimized=False)
fcast1 = fit1.forecast(3).rename(r"$\alpha=0.2$")
print(f"alpha=0.2: {fcast1.values}.")

fit2 = SimpleExpSmoothing(realty_data,
initialization_method="heuristic").fit(smoothing_level=0.
6, optimized=False)
fcast2 = fit2.forecast(3).rename(r"$\alpha=0.6$")
print(f"alpha=0.6: {fcast2.values}.")

fit3 = SimpleExpSmoothing(realty_data,

```

```
initialization_method="estimated").fit()
fcast3 = fit3.forecast(3).rename(r"\alpha=%s$" %
fit3.model.params["smoothing_level"])
print(f"alpha={fit3.model.params['smoothing_level']}:
{fcast3.values}.")

plt.figure(figsize=(12, 8))
plt.plot(realty_data, marker="o", color="black")

plt.plot(fit1.fittedvalues, marker="o", color="blue")
(line1,) = plt.plot(fcast1, marker="o", color="blue")

plt.plot(fit2.fittedvalues, marker="o", color="red")
(line2,) = plt.plot(fcast2, marker="o", color="red")

plt.plot(fit3.fittedvalues, marker="o", color="green")
(line3,) = plt.plot(fcast3, marker="o", color="green")
plt.legend([line1, line2, line3], [fcast1.name,
fcast2.name, fcast3.name])

plt.show()
```

## Список літератури

1. Кокс Д. Ф., Браун Р. В. Информация и риск в маркетинге. [Текст] - М.: Финстатистформ, 1993. ISBN 5-03-003320-3
2. Бахвалов, Н. С. Численные методы [Текст] : учеб. пособие для физ.-мат. специальностей вузов / Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков ; под общ. ред. Н. И. Тихонова. – 2-е изд. – М. : Физматлит : Лаб. базовых знаний ; СПб. : Нев. диалект, 2002. – 630 с. : ил. ; 25 см. – (Технический университет. Математика). – Библиогр.: с. 622–626. – Предм. указ.: с. 627–630. – 30000 экз. – ISBN 5-93208-043-4 (в пер.).
3. Демиденко Е.З. Оценки параметров в нелинейной регрессии. Серия «Математические методы в экономике и международных отношениях», Вып. «Проблемы эконометрического моделирования». - М.: ИМЭМО АН СССР, 1972.
4. Дженкинс Г., Ватте Д. Спектральный анализ и его приложения. Вып. 1 и 2. - М.: Мир, 1971, 1972.
5. Лукашин Ю. П., АДАПТИВНЫЕ МЕТОДЫ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ. Пункты 1-2. – «финансы и статистика», 2003.
6. Кендалл М. Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976.
7. Bailey M.G. « Prediction of an autoregressive variables subject both to disturbances and to errors of observation //J. Amer.Statist. Ass. - 1965. Vol. 60. - PP. 164-181.
8. Batty M. Monitoring an exponential smoothing forecasting system // Oper. Res. Quart. - 1969. - Vol. 20. - № 3.
9. Bates IM., Granger C.W.I. The combination of forecasts //Oper. Res. Quart. - 1969. - Vol. 20. - № 4.
10. Дрейпер Н., Сміт Р. Прикладний регресійний аналіз. - М: Статистика, 1973.
11. Кейн Е. Економічна статистика та економетрія. Вип. 1 і 2. - М: Статистика, 1977.
12. Кільдішев Г.С, Френкель А.А. Аналіз тимчасових рядів та прогнозування. - М: Статистика, 1973.
13. Клеандров Д.І., Френкель А.А. Прогнозування економічних показників за допомогою методу простого експоненціального згладжування // Статистичний аналіз економічних часових рядів та прогнозування. - М: Наука, 1973.
14. Brown R.G., Meyer R.F. The fundamental theorem of exponential smoothing//Oper. Res. - 1961. - Vol.9. - № 5.