

**МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ
ГЕТЕРОСКЕДАСТИЧНИХ ПРОЦЕСІВ**

П.І. БІДЮК

Розглядається методика побудови математичних моделей гетероскедастичних процесів і її застосування до опису динаміки часових рядів. Запропоновано спрощений тест на гетероскедастичність та алгоритм врахування імпульсних складових ряду, які суттєво перевищують його середнє значення. Побудовано функції прогнозу дисперсії як міри ризику на основі розв'язку рівнянь. Наведено приклади прогнозування реальних рядів.

ВСТУП

Гетероскедастичні процеси, тобто процеси із змінною в часі дисперсією, — поширений клас фінансово-економічних процесів, особливо в нестійкій перехідній економіці. За визначенням такі процеси належать до класу слабо нестационарних процесів, а тому опис їх динаміки має деякі особливості у порівнянні із стаціонарними. Одним із популярних підходів до моделювання нестационарних процесів є метод групового врахування аргументів [1], який був успішно застосований до побудови моделей низки технічних, екологічних та економічних процесів. Частково питання побудови моделей нестационарних процесів розглянуто в роботах [2 – 7], однак моделювання гетероскедастичних процесів потребує докладного розгляду.

Метою даної роботи є розробка спрощеного тесту на гетероскедастичність, універсальної методики моделювання гетероскедастичних процесів, а також побудова моделей деяких реальних фінансово-економічних процесів та функцій прогнозування на їх основі.

1. ТЕСТУВАННЯ НА НАЯВНІСТЬ ГЕТЕРОСКЕДАСТИЧНОСТІ

При моделюванні гетероскедастичних процесів спочатку виконують перевірку на наявність гетероскедастичності. В літературі наведено кілька тестів на гетероскедастичність [6, 8, 9], які принципово не відрізняються один від одного, але потребують додаткових і не завжди виправданих обчислень. Найбільш зручним є такий.

Спрощений тест на гетероскедастичність складається з кроків: 1) оцінити авторегресію $y(k) = a_0 + a_0 y(k-1) + \varepsilon(k)$ або більш високого порядку, наприклад, другого чи третього; 2) побудувати ряд $\{\varepsilon^2(k)\}$, скориставшись залишками від оцінювання попередньої моделі; 3) оцінити регресію $\varepsilon^2(k) = a_0 + \alpha_1 [0,4\varepsilon(k-1) + 0,3\varepsilon(k-2) + 0,2\varepsilon(k-3) + 0,1\varepsilon(k-4)]$; 4) якщо коефіцієнт α_1 відмінний від нуля в статистичному сенсі, тобто є значимим, то отримана модель для $\varepsilon^2(k)$ описує гетероскедастичний процес.

Оскільки в цій моделі оцінюються тільки коефіцієнти α_1 і a_0 , а всі інші відомі (0,4; 0,3; 0,2; 0,1), то для визначення відмінності від нуля коефіцієнта α_1 можна застосувати теорію перевірки гіпотез, тобто t -статистику для цього коефіцієнта.

Привабливість даного тесту полягає в простих обчисленнях і можливості застосування тієї ж теорії перевірки гіпотез, що використовується при аналізі лінійних моделей.

Сформулюємо постановку задачі оцінювання та прогнозування для авторегресійних умовно гетероскедастичних моделей АРКС (p, q) на основі максимізації функції правдоподібності.

Дана послідовність значень часового ряду $\{y(1), y(2), \dots, y(N)\}$. Для класу лінійних та псевдолінійних моделей АР (p) та АРУГ

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \varepsilon_1(k),$$

$$\varepsilon_1^2(k) = \alpha_0 + \alpha_1 [\beta_1 \varepsilon_1^2(k-1) + \beta_2 \varepsilon_1^2(k-2) + \dots + \beta_m \varepsilon_1^2(k-m)] + \varepsilon_2(k),$$

де $\varepsilon_1(k)$ – послідовність випадкових величин із змінною дисперсією; $\varepsilon_2(k)$ – послідовність випадкових величин, яким властиве

$$E[\varepsilon_2(k)] = 0, \quad E[\varepsilon_2(k)\varepsilon_2(j)] = \begin{cases} \sigma^2, & k = j, \\ 0, & k \neq j. \end{cases}$$

Необхідно знайти:

1) оцінку вектора параметрів $\theta_1 = [a_0 \ a_1 \dots \ a_p]^T$, мінімізуючи критерій

$$J = \sum_{k=1}^N [y(k) - \hat{y}(k)]^2, \quad \text{де } \hat{y}(k) = \hat{a}_0 + \sum_{i=1}^p \hat{a}_i y(k-i) + \varepsilon_1(k);$$

2) оцінки вектора $\theta_2 = [\alpha_0 \ \alpha_1 \ \beta_1 \dots \ \beta_m]^T$ із умови максимізації функції правдоподібності $L(\theta_2) = \log(f_{y_1}(y_{1m}; \theta_2)) \prod_{k=2}^N (f_{y_k|y_{k-1}}(y_{km}|y_{(k-1)m}; \theta_2))$;

3) оцінку прогнозу дисперсії на основі оціненої моделі $\hat{\varepsilon}^2(k+s|k) = E_k[\varepsilon^2(k)]$, $s=1, 2, \dots$, використовуючи всю наявну інформацію на момент часу k , із умови $J_1 = E[\hat{\varepsilon}^2(k+s), \hat{\varepsilon}^2(k+s)] \rightarrow \min$.

2. МЕТОДИКА ПОБУДОВИ МОДЕЛЕЙ ГЕТЕРОСКЕДАСТИЧНИХ ПРОЦЕСІВ

Для прикладу розглянемо спочатку узагальнену авторегресійну модель гетероскедастичного процесу (УАРУГ). При використанні такої моделі умовна дисперсія процесу описується за допомогою моделі авторегресії з ковзним середнім (АРКС). Нехай похибки моделі

$$\varepsilon(k) = v(k) \sqrt{h(k)}, \quad (1)$$

де $v(k)$ — процес білого шуму з одиничною (для простоти) дисперсією $\sigma_v^2 = 1$;

$$h(k) = a_0 + \sum_{i=1}^q \alpha_i \varepsilon^2(k-i) + \sum_{i=1}^p \beta_i h(k-i), \quad (2)$$

де $h(x)$ — умовна дисперсія.

Оскільки процес $\{v(k)\}$ визначено в даному випадку як білий шум, то умовне і безумовне середнє процесу будуть дорівнювати нулю.

Математичне сподівання для $\varepsilon(k)$

$$E[\varepsilon(k)] = E\{v(k) \sqrt{h(k)}\} = 0.$$

Важливим моментом є те, що умовна дисперсія процесу $\{\varepsilon(k)\}$ залежна від часу $E_{k-1}[\varepsilon^2(k)] = h(k)$, оскільки $E_{k-1}[v^2(k)] = 1$. Для того щоб ця умовна дисперсія була скінченною, необхідно щоб корені характеристичного рівняння, записаного для (2), знаходились всередині кола одиничного радіуса. Таким чином, основною відмінною властивістю моделі УАРУГ є те, що умовна дисперсія збурень, які діють на процес $\{y(k)\}$, — це процес авторегресії з ковзним середнім.

Припустимо, що $\{y(k)\}$ — процес авторегресії з ковзним середнім. При побудові моделі процесу в даному випадку можливі варіанти:

1) якщо вдається побудувати адекватну модель АРКС, то похибки моделі будуть мати властивості білого шуму;

2) якщо не вдається побудувати таку модель, то, використовуючи автокореляційну функцію для квадратів залишків, необхідно побудувати модель УАРУГ, яка дозволить виконати аналіз поведінки дисперсії процесу (корелограма процесу $\{\varepsilon^2(k)\}$ дасть можливість визначити присутність гетероскедастичності).

Оскільки $E_{k-1}[\varepsilon(k)] = \sqrt{h(k)}$, то рівняння (2) можна записати у вигляді

$$E_{k-1}[\varepsilon^2(k)] = a_0 + \sum_{i=1}^q \alpha_i \varepsilon^2(k-i) + \sum_{i=1}^p \beta_i h(k-i). \quad (3)$$

За своєю структурою це рівняння подібне до рівняння АРКС (q, p) для послідовності $\{\varepsilon^2(k)\}$. Таким чином, методику побудови моделі гетероскедастичного процесу можна узагальнити.

Крок 1. При необхідності зробити попередню обробку експериментальних даних (нормування, логарифмування, заповнення пропусків даних) і застосувати до них тести на гетероскедастичність. Якщо процес містить тренд, то перед побудовою моделі необхідно його видалити. Досить часто візуальний аналіз даних дозволяє отримати суттєву інформацію щодо присутності гетероскедастичності. Разом із візуальним аналізом необхідно розглядати параметри описової статистики, які полегшують визначення структури моделі.

Крок 2. Користуючись АКФ та ЧАКФ для експериментальних даних, побудувати модель АР(p) або АРКС(q, p) для процесу $\{y(k)\}$ та обчислити ряд з квадратів залишків $\{\hat{\varepsilon}^2(k)\}$, де $\hat{\varepsilon}(k) = e(k)$. Обчислити вибірккову дисперсію $\hat{\sigma}_{\varepsilon}^2$ збурення $\hat{\varepsilon}(k)$: $\hat{\sigma}_{\varepsilon}^2 = \frac{1}{N-1} \sum_{k=1}^N \hat{\varepsilon}^2(k)$, де N — число залишків після побудови моделі АР чи АРКС.

Крок 3. Обчислити і побудувати графік вибіркової автокореляційної функції для квадратів залишків

$$\rho(s) = \frac{\sum_{k=s+1}^N [\hat{\varepsilon}^2(k) - \hat{\sigma}_{\varepsilon}^2][\hat{\varepsilon}^2(k-s) - \hat{\sigma}_{\varepsilon}^2]}{\sum_{k=1}^N [\hat{\varepsilon}^2(k) - \hat{\sigma}_{\varepsilon}^2]} \quad (4)$$

Якщо існують такі значення $\rho(s)$, які відрізняються від нуля із статистичної точки зору, то це свідчить про присутність процесу АРУГ або УАРУГ. Для того щоб переконатись у присутності гетероскедастичності, використовують Q -статистику Люнга-Бокса, яка обчислюється за виразом

$$[7]: Q = N(N+2) \sum_{i=1}^n \rho(i)/(N-i), \quad \text{де } n = N/4 \text{ (емпіричне значення).}$$

Якщо значення $\hat{\varepsilon}^2(k)$ некорельовані, то Q -статистика повинна мати розподіл χ^2 з n ступенями свободи.

Крок 4. Побудувати модель УАРУГ (або іншу модифікацію)

$$h(k) = a_0 + \sum_{i=1}^q \alpha_i \varepsilon^2(k-i) + \sum_{i=1}^p \beta_i h(k-i), \quad (5)$$

використовуючи ряд значень $\hat{\varepsilon}^2(k)$. Якщо в цій моделі хоча б один із коефіцієнтів $\alpha_i, i=1, \dots, q$ є значимим, то процес дійсно гетероскедастичний. Оскільки модель (5) описує залишки моделі з деяким наближенням, то в загальному випадку доцільно продовжити процес уточнення моделей, що описують вихідний процес в цілому. Тобто, можна уточнити як початкову модель АР(p) чи АРКС(p, q), а також модель типу (2). Це робиться на наступному кроці.

Крок 5. Скористатись моделлю (5), для того щоб отримати дійсні значення залишків, які описуються цією моделлю, тобто згенерувати ряд

$\{\hat{\varepsilon}_1(k)\}$. Згенерувати ще один ряд $\{y_1(k)\}$, де $y_1(k) = y(k) - \hat{\varepsilon}_1(k)$. За допомогою отриманого ряду побудувати уточнену модель процесу типу АР(p) чи АРКС(p, q). При необхідності процес уточнення моделей може бути продовжений.

3. АЛГОРИТМ ВРАХУВАННЯ В МОДЕЛІ АРУГ ІМПУЛЬСНИХ АДИТИВНИХ СКЛАДОВИХ ЧАСОВОГО РЯДУ

Моделі АРУГ дозволяють врахувати ефекти, присутні в часових рядах, які не можуть бути описані математично за допомогою інших підходів до моделювання. Так, узагальнена модель АРУГ використовується для опису даних, що мають надлишковий куртозис. Стандартну модель УАРУГ можна використати також для опису похибок моделювання, які мають t -розподіл (модель УАРУГ- t). Однак ще існують інші проблеми у досягненні високого ступеня адекватності моделей такого типу, наприклад, випадки, коли ряд містить адитивні складові (імпульси), що суттєво перевищують поточне середнє значення, а також ближні сусідні значення. Така проблема існує при побудові моделей для цін акцій на біржах, а також при моделюванні доходів від цінних паперів. Тому ми пропонуємо підхід до побудови моделей УАРУГ, які враховують вказаний ефект, що дозволяє суттєво підвищити адекватність моделі.

Розглянемо часовий ряд $\{y(k)\}$, $k=1,2,\dots,N$, який може бути описаний з деяким ступенем адекватності моделлю АРКС(p, q), тобто

$$y(k) = \sum_{i=1}^p a_i y(k-i) + \sum_{j=0}^q b_j \varepsilon(k-j), \quad (6)$$

де $\varepsilon(k)$ – випадкова послідовність типу білого шуму з нульовим середнім та дисперсією σ_ε^2 . В моделі (6) відсутня константа a_0 , оскільки припускається, що вона будується на основі вимірів $y(k) - \mu_y$, де μ_y — середнє значення ряду. Припустимо також, що рівняння (6) має збіжний розв’язок, тобто корені його характеристичного рівняння знаходяться в колі одиничного радіуса. Рівняння (6) можна записати також із застосуванням оператора зсуву L як

$$(1 - a_1 L - a_2 L^2 - \dots - a_p L^p) y(k) = (1 + b_1 L + b_2 L^2 + \dots + b_q L^q) \varepsilon(k),$$

де L — оператор затримки в часі; $Ly(k) = y(k-1)$.

Нехай модель, яка враховує значні адитивні імпульсні складові часового ряду, має вигляд

$$z(k) = y(k) + \beta u(k, l), \quad (7)$$

де

$$u(k, l) = \begin{cases} 1, & k=l, \\ 0, & k \neq l. \end{cases}$$

Тобто реальні спостереження описуються рівнянням (7), а не (6). Розглянемо приклад, коли часовий ряд містить одну адитивну складову (значний імпульс), який суттєво перевищує поточне середнє значення.

При оцінюванні моделі АРКС (p, q) на основі такого ряду залишки моделі опишемо так:

$$\hat{\varepsilon}(k) = \frac{A(L)}{B(L)} y(k) = D(L) y(k), \quad (8)$$

де $D(L) = A(L)/B(L) = (1 - d_1 L - d_2 L^2 - \dots)$. Значення порядку авторегресії p та ковзного середнього q визначимо в процесі побудови моделі (6). Якщо ряд містить значний імпульс, то рівняння (8) приймає вигляд

$$\hat{\varepsilon}(k) = \varepsilon(k) + \beta D(L) u(k, l), \quad (9)$$

і його можна розглядати як модель регресії для $\hat{\varepsilon}(k)$

$$\hat{\varepsilon}(k) = \beta \omega(k) + \varepsilon(k), \quad (10)$$

де

$$\omega(k) = \begin{cases} 0, & k < l, \\ 1, & k = l, \end{cases} \text{ і } \omega(k+j) = -D(j), \quad k > l, \quad j = 1, 2, \dots$$

Тепер вплив суттєвого імпульсного значення на модель в момент $k = l$ запишемо як

$$\hat{\beta}(l) = \frac{\sum_{k=l}^N \hat{\varepsilon}(k) \omega(k)}{\sum_{k=l}^N \omega^2(k)}. \quad (11)$$

Функцію $\hat{\beta}(l)$ представляють у вигляді, зручному для тестування її значимості. Для цього необхідно визначити оцінку дисперсії залишків, отриманих при побудові моделі (6). При цьому бажано, щоб оцінка $\hat{\beta}(l)$ не була суттєво зміщеною внаслідок присутності імпульсів у вихідному ряді значень. Оцінка дисперсії залишків може бути отримана різними способами. Один з них полягає в тому, що дисперсія обчислюється для залишків моделі, з яких видаляється спостереження (імпульс) при $k = l$. Якщо позначити цю дисперсію $\hat{\sigma}_e^2$, то стандартний статистичний параметр (статистика) має вигляд

$$\hat{\tau} = \frac{\hat{\beta}(l) / \hat{\sigma}_e}{\left[\sum_{k=l}^N \omega^2(k) \right]^{1/2}}. \quad (12)$$

Вплив імпульсного значення ряду вважається значним, якщо $\hat{\tau} > c$, де c — деяке порогове значення. Оскільки в асимптотиці $\hat{\tau}$ має нормальний

розподіл, то порогове значення c можна визначити з таблиць для t -розподілу. На основі емпіричних досліджень пропонується використовувати для цього величину $c=4$. Якщо $\hat{t} > c$, можна скоригувати спостереження $z(k)$ в момент $k=l$, і таким чином перейти до ряду $y(k)$ (без імпульсного значення), використовуючи $\hat{\beta}$, отриману за допомогою (11), тобто $y(k) = z(k) - \hat{\beta}u(k, l)$.

Розглянутий підхід до врахування імпульсних значень часового ряду даних поширюється на випадок наявності множини імпульсів, тобто процедура повторюється для кожного імпульсу доти, поки \hat{t} -статистика не стане незначимою. На останньому кроці параметри моделі оцінюють заново для всіх спостережень.

Розглянемо застосування запропонованої вище методики до побудови моделей гетероскедастичних процесів. Нехай $p(k)$ — ціна акцій на біржі або індекс біржових цін. Сформуємо ряд з логарифмів перших різниць як

$$d(k) = \log[p(k)] - \log[p(k-1)]$$

і запишемо можливу структуру узагальненої авторегресійної умовно гетероскедастичної моделі у вигляді

$$d(k) = \eta(k)h^{1/2}(k), \quad (13)$$

$$h(k) = \alpha_0 + \alpha_1 d^2(k-1) + \gamma_1 h(k-1), \quad (14)$$

де $h^{1/2}(k)$ — стандартне відхилення для умовної дисперсії, обчисленої на основі ряду даних. Якщо зробити припущення, що $\eta(k)$ — некорельована нормально розподілена послідовність, то рівняння (13), (14) представляють собою модель УАРУГ. Але якщо $\eta(k)$ має умовний t -розподіл, то модель (13), (14) називають УАРУГ- t . Для покращання опису ряду з імпульсними значеннями скористаємось стандартною моделлю УАРУГ, але будемо коригувати її із врахуванням присутності імпульсних значень в ряді $\eta(k)$.

Рівняння (14) описує умовну мінливість величин $d(k)$, а тому ним можна скористатись для прогнозування мінливості цін, доходів і т.п. Представимо модель УАРУГ у вигляді

$$d^2(k) = \alpha_0 + (\alpha_1 + \gamma_1)d^2(k-1) + x(k) - \gamma_1 x(k-1), \quad (15)$$

де $x(k) \equiv d^2(k) - h(k)$. Рівняння (15) — це модель типу АРКС(1,1) для $d^2(k)$. Зазначимо, що ряд $x(k)$ можна розглядати як гетерогенний ряд залишків. Тепер алгоритм побудови моделі із врахуванням присутності імпульсних значень можна представити у вигляді таких кроків:

Крок 1. Обчислити оцінки параметрів рівнянь (13), (14) методом максимальної правдоподібності. В результаті отримаємо оцінки $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\gamma}_1$, а також оцінки рядів $\hat{h}(k)$, $\hat{\eta}(k)$. Це дозволить знайти ряд $\hat{x}(k) = d^2(k) - \hat{h}(k)$,

який використовується на наступному кроці при визначенні присутності імпульсних значень.

Крок 2. З рівняння (8) випливає, що $\hat{D}(L) = (1 - (\hat{\alpha}_1 + \hat{\gamma}_1)L) / (1 - \hat{\gamma}_1L)$. Тепер для $\forall k=l$ необхідно оцінити рівняння регресії (10), використовуючи оцінки залишків $\hat{z}(k)$, та обчислити $\hat{\beta}_z(l)$. Незважаючи на те, що ряд $\hat{x}(k)$ гетерогенний, необхідно обчислити безумовну дисперсію для $\hat{x}(k)$ без врахування імпульсних значень ряду. Тобто дисперсія похибки моделі обчислюється при нульовому значенні спостереження та $k=l$. Для того щоб врахувати зміну в часі дисперсії величини $\hat{x}(k)$, її необхідно обчислити для вибраного рухомого вікна, а не для вибірки в цілому. Статистики $\hat{\tau}$, які визначаються виразом (12), обчислюються для $\hat{x}(k)$ і порівнюються із значенням $c=4$.

Крок 3. Спостереження $\hat{x}(k)$ при $k=l$ та найбільшому значенні статистики $\hat{\tau}$, яке повинне перевищувати $c=4$, замінюється на $\hat{x}^*(k)$ (так, як це показано в рівнянні (7)) з використанням вагового коефіцієнта $\hat{\beta}(l)$ (вираз (11)).

Крок 4. За допомогою ряду $\hat{x}^*(k)$ та $\hat{h}(k)$ побудувати ряд $d^{*2}(k)$, тобто обчислити його за виразом $d^{*2}(k) = \hat{x}^*(k) + \hat{h}(k)$ для моменту часу $k=l$. Скорегувати ряд даних відносно імпульсних значень

$$d^*(k) = \begin{cases} d(k), & k \neq l, \\ \text{sign}[d(k)][d^{*2}(k)]^{1/2}, & k = l. \end{cases} \quad (16)$$

Наведений вираз показує, що $\text{sign}[d^2(k)] = \text{sign}[d(k)]$ при $k=l$. З цього випливає: для членів ряду, які корегуються у відповідності до присутності імпульсного значення, знак не змінюється.

Крок 5. Після обчислення значень ряду $d^*(k)$ перейти на крок 1 і повторювати всю наведену процедуру доти, доки $\hat{\tau} > 4$, тобто доки не будуть оброблені всі члени ряду, які мають суттєві імпульсні значення. Після обробки даних отримаємо остаточні значення оцінок $\hat{\alpha}_0^*$, $\hat{\alpha}_1^*$, $\hat{\gamma}_1^*$ і ряди даних $d^*(k)$, $\hat{h}^*(k)$. За допомогою цих оцінок параметрів та рядів обчислюється прогноз на один крок, тобто $\hat{h}^*(N+1)$ у відповідності до рівняння (14).

Якщо $d(k)$ в рівнянні (13) представляє собою корельований процес, то перед використанням запропонованого п'ятикрокового алгоритму необхідно відфільтрувати детерміновану динаміку процесу за допомогою моделі АР або АРКС. В результаті отримаємо залишки моделі $\varepsilon(k)$ так само, як це було раніше при побудові моделі АРУГ. Після цього отримані залишки поступають на вхід алгоритму. Однак необхідно пам'ятати, що коефіцієнт корекції β в рівнянні (7) необхідно застосовувати до початкових даних $d(k)$, а не до залишків $\varepsilon(k)$, оскільки структура моделі АРКС може бути іншою

після застосування корекції внаслідок присутності суттєвих імпульсних значень.

Приклад 3.1. Розглянемо побудову моделі УАРУГ для ряду $d(k)$ та моделі УАРУГ для скоригованого відносно імпульсних значень ряду $d^*(k)$. Для прикладу візьмемо дані, які характеризують щотижневий прибуток російської біржі сільгосппродуктів за 1996 – 1999 рр. Процедура обробки статистичних даних містить оцінювання параметрів математичних моделей на основі щотижневих чотирирічних даних, які використовуються для обчислення оцінки однокрокового прогнозу для дисперсії $h(k)$. Потім з вибірки даних видаляється перше значення (за перший тиждень всього періоду) і одне значення додається в кінці вибірки (наступний тиждень). Знову оцінюються параметри моделей і обчислюється однокроковий прогноз. Таким чином запропонований метод корегування впливу імпульсних значень застосовується послідовно до всіх значень вибірки.

На основі даного часового ряду оцінено моделі $AR(p)$, потім УАРУГ (1,1) і модель для прибутку, що має таку структуру:

$$d(k) = \mu + a_1 d(k-1) + a_2 d(k-2) + \varepsilon(k),$$

$$\varepsilon(k) = \eta(k) h^{1/2}(k), \quad \eta(k) \sim N(0,1),$$

$$h(k) = \alpha_0 + \alpha_1 \varepsilon^2(k-1) + \gamma_1 h(k-1).$$

Результати оцінювання наведено в табл. 1. Параметри моделей оцінено методом максимальної правдоподібності. Діагностичні статистичні параметри Q_1, Q_2 , наведені в двох останніх стовпчиках табл. 1, свідчать про те, що моделі мають достатню ступінь адекватності. Параметри $Q_1(10)$ і $Q_2(10)$ — це статистики Льюнга-Бокса, які обчислені для визначення ступеня кореляції залишків моделі. Критичне значення для Q -статистики визначається за допомогою таблиць для розподілу χ^2 . На наступному кроці до оцінених моделей застосовується метод корекції моделі із врахуванням присутності суттєвих імпульсних значень ряду.

Таблиця 1. Результати оцінювання моделей $AR(p)$ та УАРУГ(1,1)

Вибірка даних, рік	Параметри моделі						Діагностика	
	$\hat{\mu}$	\hat{a}_1	\hat{a}_2	α_0	α_1	γ_1	$Q_1(10)$	$Q_2(10)$
1996–1999	0,0034 (1,97)	0,0017 (2,05)	0,179 (2,73)	0,0002 (2,18)	0,274 (4,61)	0,468 (4,97)	5,48	5,17
1997–2000	0,0015 (2,36)	0,0011 (1,99)	0,196 (2,57)	0,0003 (3,19)	0,385 (3,27)	0,652 (4,09)	6,85	7,33

В табл. 2 наведено результати оцінювання тих же моделей, але вже із врахуванням корекції внаслідок присутності суттєвих імпульсних значень, тобто модель побудована для $d^*(k)$.

Таблиця 2. Результати оцінювання моделей $AR(p)$ та $ARUG(1,1)$ після корекції

Вибірка даних, рік	Параметри моделі						Діагностика	
	$\hat{\mu}$	\hat{a}_1	\hat{a}_2	α_0	α_1	γ_1	$Q_1(10)$	$Q_2(10)$
1996–1999	0,0043 (2,45)	0,0021 (2,39)	0,353 (3,19)	0,0002 (2,18)	0,118 (4,02)	0,694 (5,37)	5,41	4,76
1997–2000	0,0012 (2,44)	0,0009 (1,78)	0,105 (2,69)	0,0002 (3,19)	0,157 (3,06)	0,774 (4,51)	7,49	5,28

У порівнянні з табл. 1 оцінки параметрів α_1 і γ_1 змінилися суттєво. Наприклад, для періоду даних 1996 – 1999 рр. значення α_1 зменшилося від 0,274 до 0,118, а γ_1 збільшилося від 0,468 до 0,694. Однак сума коефіцієнтів $\alpha_1 + \gamma_1$ суттєво не змінилася, що свідчить про те, що загалом динаміка зміни доходів є робастною по відношенню до імпульсних значень. На основі отриманих даних моделювання можна зробити висновок, що застосування корекції даних приводить до більш стійких оцінок параметрів.

Приклад 3.2. Аналіз і моделювання утворення цін на продукцію виробничої фірми (рис. 1).

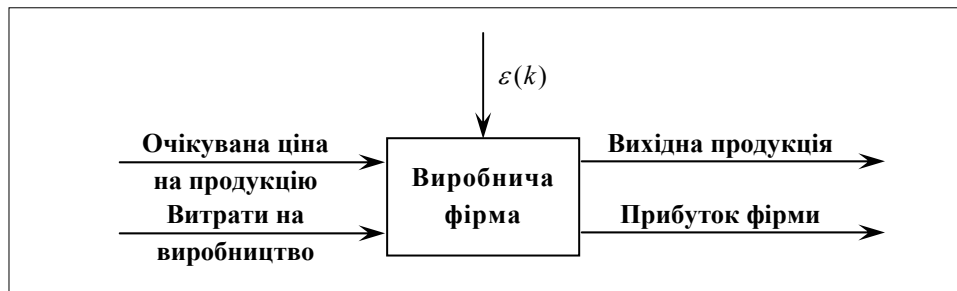


Рис. 1. Вхідні та вихідні змінні виробничої фірми

Нехай $q(k)$ – прибуток фірми (дебітора) в момент часу k ; $p^e(k)$ — очікувана ціна продукції фірми в момент k , отримана на основі інформації в момент часу $(k-1)$ (тобто $p^e(k) = E_{k-1}[p(k)]$); $h(k)$ — очікувана дисперсія ціни на продукцію фірми в момент k на основі інформації в момент часу $(k-1)$; $pp(k-1)$ – затрати на виробництво одиниці продукції в момент $(k-1)$; $p(k-1)$ — об'єм виробництва продукції в момент часу $(k-1)$; $\varepsilon_1(k)$ — випадкові збурення прибутку фірми.

Для опису прибутку вибрана структура моделі

$$q(k) = a_0 + a_1 p^e(k) + a_2 h(k) + a_3 pp(k-1) + a_4 p(k-1) + a_5 q(k-4) + \varepsilon_1(k). \quad (17)$$

Дані вимірюються щоквартально. Для обчислення $p^e(k)$ і $h(k)$ будуються окремі моделі: $pp(k-1)$ — затрати на виробництво одиниці продукції

за попередній квартал; $q(k-4)$ — береться з запізненням 4, щоб врахувати той же квартал попереднього року. Що цікавого в цій моделі? Це — вплив умовної дисперсії ціни продукції на прибуток фірми. Змінна $p^e(k)$ — оцінюється на основі ціни в попередньому кварталі.

Якщо ціна змінюється достатньо швидко, то виробник, уникаючи ризику, намагається знизити об'єм виробництва. Для прогнозу рівня ціни на продукцію використано модель четвертого порядку

$$(1 - \beta_1 L - \beta_2 L^2 - \beta_3 L^3 - \beta_4 L^4) p(k) = \beta_0 + \varepsilon_2(k). \quad (18)$$

$$(1 - 0,5111L - 0,129L^2 - 0,130L^3 - 0,138L^4) p(k) = 1,632 + \varepsilon_2(k). \quad (19)$$

Тепер беруть $p^e(k) = p(k)$ для того, щоб підставити у (17). Передбачається, що динаміка ціни — гетероскедастичний процес:

$$h(k) = 1,353 + 0,162\varepsilon_2^2(k-1) + 0,591h(k-1). \quad (20)$$

Значення, обчислені в (19) і (20) підставляються у (17) для оцінки рівня цін на продукцію. Коефіцієнти в (19) і (20) відносяться до прибутку фірми у виробництві бройлерів. Рівняння для прибутку:

$$q(k) = 2,767 p(k) + 0,521 h(k) + 4,325 pp(k-1) + 1,887 p(k-1) + 0,603 q(k-4) + \varepsilon(k).$$

Приклад 3.3. Ціни акцій на біржі. Розглянемо ряд $\varepsilon^2(k)$ для моделі, побудованої у попередньому розділі (приклад 1) АР (АР(1,5,8)) (коливання цін акцій на біржі, рис.2; часткова АКФ на рис. 3).

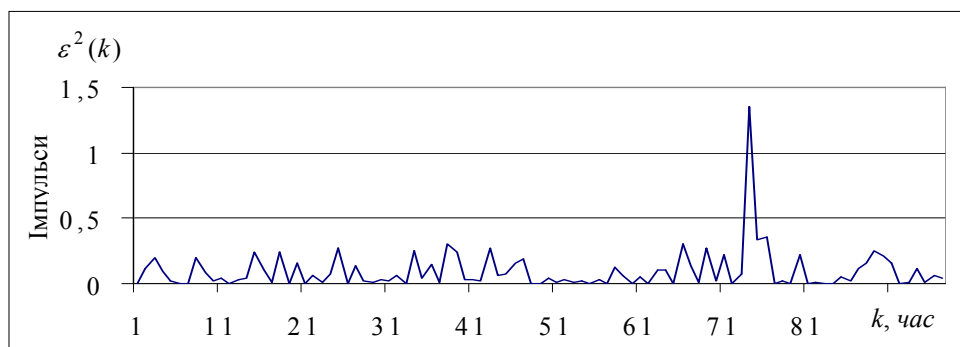


Рис. 2. Автокореляційна функція ряду

В результаті дослідження АКФ і ЧАКФ встановлено:

1. Значення коефіцієнтів АКФ: $\rho_1 = 0,124$; $\rho_2 = 0,107$; $\rho_4 = -0,066$; $\rho_5 = 0,103$; $\rho_{10} = -0,069$.

2. Значення коефіцієнтів ЧАКФ: $\Phi_{1,1} = 0,124$; $\Phi_{2,2} = 0,093$; $\Phi_{4,4} = -0,072$; $\Phi_{5,5} = 0,127$; $\Phi_{10,10} = -0,91$.

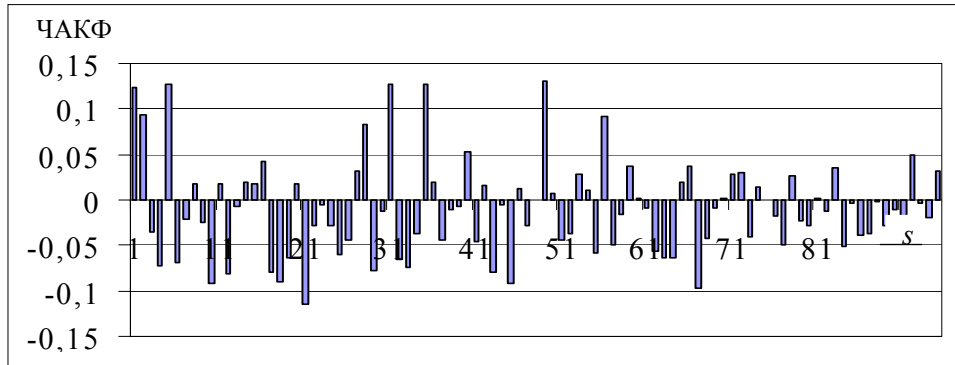


Рис. 3. Часткова автокореляційна функція ряду

3. Для математичного опису процесу необхідно розглянути моделі $AR(1)$, $AR(2)$, $AR(3)$, $AR(4)$. Найбільш вірогідні номери запізнювань, які входять до складу моделі: 1, 2, 4, 5, 10. У табл. 3 наведено варіанти оцінювання декількох можливих структур регресійної моделі.

Таблиця 3. Варіанти оцінювання дисперсії вартості акцій однієї з компаній, що входять до числа провідних на Нью-Йоркській фондовій біржі за 1996 р.

	(1)	(1, 2)	(1, 2, 4)	(1, 2, 4, 5)	(1, 2, 4, 5, 10)
a_0	0,0851 (4,4618)	0,0769 (3,6166)	0,0822 (3,4885)	0,0714 (2,7802)	0,0814 (2,8342)
a_1	0,1242 (1,2149)	0,1138 (1,0956)	0,1110 (1,0549)	0,1187 (1,1231)	0,1145 (1,0484)
a_2		0,0927 (0,8945)	0,1056 (0,9999)	0,1123 (1,0581)	0,1192 (1,0871)
a_3			-0,0762 (-0,7276)	-0,0960 (-0,9054)	-0,1016 (-0,9274)
a_4				0,1251 (1,1845)	0,1247 (1,1385)
a_5					-0,0856 (-0,7837)
RSS	2,4041	2,3826	2,3560	2,3106	2,2628
AIC	-0,8076	-0,7846	-0,7517	-0,7377	-0,6735
BSC	-0,7542	-0,7040	-0,6428	-0,6006	-0,5034
DW	2,0196	1,9899	1,9712	1,9771	1,9738
F	0,0157	0,0248	0,0311	0,0497	0,0566
R^2	0,0155	0,0242	0,0307	0,0474	0,0536

З табл. 3 видно, що найкраща модель, яка описує дисперсію з обчислених варіантів — це $AR(1, 2)$, оскільки для неї $DW = 1,9899$, $AIC = -0,7846$. Таким чином, модель процесу має вигляд

$$\varepsilon^2(k) = 0,0769 + 0,1138\varepsilon^2(k-1) + 0,0927\varepsilon^2(k-2).$$

Значення $RSS = 2,3826$ і $R^2 = 0,0242$ свідчать, що модель можна покращити за рахунок введення ковзного середнього. Варіанти оцінювання моделей з ковзним середнім наведені в табл. 4.

Таблиця 4. Варіанти оцінювання моделей з ковзним середнім

	(1)	(1, 2)	(1, 2, 4)	(1, 2, 4, 5)	(1, 2, 4, 5, 10)
a_0	0,0851 (6,18E+14)	0,0769 (1,24E+16)	0,0822 (1,80E+15)	0,0714 (7,47E+14)	0,0814 (6,23E+14)
a_1	0,1242 (8,94E+13)	0,1138 (7,93E+13)	0,1110 (2,12E+14)	0,1187 (1,21E+14)	0,1145 (9,07E+13)
a_2		0,0927 (4,16E+14)	0,1056 (1,15E+15)	0,1123 (5,60E+14)	0,1192 (4,34E+14)
a_3			-0,0762 (-1,18E+15)	-0,0960 (-5,86E+14)	-0,1016 (-4,38E+14)
a_4				0,1251 (7,10E+14)	0,1247 (5,07E+14)
a_5					-0,0856 (-3,71E+14)
a_6	-1,34E-14 (-9,6131)	-1,35E-14 (-9,3986)	-4,60E-15 (-8,7229)	-8,91E-15 (-8,9278)	-1,12E-14 (-8,6915)
RSS	6,69E-30	3,83E-30	8,77E-31	5,27E-30	9,71E-30
DW	0,4843	1,1944	1,2519	2,0675	2,0796
R^2	1,0000	1,0000	1,0000	1,0000	1,0000

З табл. 4 видно, що краща модель обов'язково повинна містити ковзне середнє. З обчислених варіантів моделей кращою є АРКС ((1,2,4,5), (1)), оскільки для неї $DW = 2,0675$, $R^2 = 1$. Крім того, велика роль t -статистики для оцінок коефіцієнтів свідчить про їх значимість із статистичної точки зору. Таким чином, модель, яка адекватно описує дисперсію, можемо записати так:

$$\varepsilon_1^2(k) = 0,0714 + 0,1187\varepsilon_1^2(k-1) + 0,1123\varepsilon_1^2(k-2) - 0,096\varepsilon_1^2(k-4) + 0,1257\varepsilon_1^2(k-5) + \varepsilon_2(k).$$

Приклад 3.4. Ціни акцій компанії УКРНАФТА. Розглянемо ряд $\varepsilon^2(k)$ для побудованої в попередньому розділі (приклад 2) моделі АР (1,2,3) щоденної вартості акцій компанії УКРНАФТА за 2001 р. В результаті дослідження АКФ і ЧАКФ встановлені такі значення коефіцієнтів:

1. АКФ: $\rho_1 = 0,166$; $\rho_2 = -0,006$; $\rho_3 = 0,028$; $\rho_9 = 0,038$; $\rho_{10} = 0,132$.
2. ЧАКФ: $\Phi_{1,1} = 0,166$; $\Phi_{2,2} = -0,035$; $\Phi_{3,3} = 0,036$; $\Phi_{9,9} = 0,043$; $\Phi_{10,10} = 0,122$.

Для математичного опису процесу необхідно розглянути моделі AP(1), AP(2), AP(3), AP(4). Найбільш вірогідні номери запізнювань, які входять до складу моделі: 1, 2, 3, 9, 10. У табл. 5 наведені варіанти оцінювання декількох можливих структур регресійної моделі.

Таблиця 5. Варіанти оцінювання дисперсії щоденної вартості акцій компанії УКРНАФТА за 2001 р.

	(1)	(1, 2)	(1, 2, 3)	(1, 2, 3, 9)	(1, 2, 3, 9, 10)
a_0	0,1652 (5,1101)	0,1777 (5,2338)	0,1738 (4,8391)	0,1606 (4,1112)	0,1350 (3,3577)
a_1	0,1666 (3,1342)	0,1199 (1,9116)	0,1180 (1,8648)	0,1230 (1,9169)	0,1164 (1,8261)
a_2		-0,0263 (-0,4872)	-0,0495 (-0,7817)	-0,0458 (-0,7101)	-0,0429 (-0,6701)
a_3			0,0369 (0,6801)	0,0507 (0,7907)	0,0504 (0,7913)
a_4				0,0397 (0,7376)	0,0403 (0,6384)
a_5					0,1218 (2,2372)
RSS	58,5397	57,8552	57,6733	57,1763	55,8335
AIC	1,3820	1,3818	1,3909	1,4151	1,4037
BSC	1,4098	1,4235	1,4468	1,4862	1,4892
DW	2,0848	2,0029	1,9978	2,0121	2,0110
F	0,0389	0,0146	0,0164	0,0201	0,0440
R^2	0,0374	0,0144	0,0161	0,0197	0,0421

Аналіз отриманих результатів дозволяє зробити такі висновки:

Найкраща модель, яка описує дисперсію з обчислених варіантів — AP(1, 2), оскільки для неї $DW = 2,0029$, $R^2 = 0,0144$, $AIC = 1,4235$ (табл. 5). Таким чином, модель, що адекватно описує процес, запишемо як

$$\varepsilon_1^2(k) = 0,1777 + 0,1199\varepsilon_1^2(k-1) - 0,0263\varepsilon_1^2(k-2).$$

Значення $RSS = 57,8552$ показує, що модель можна поліпшити за рахунок введення ковзного середнього. Оцінки параметрів моделей з ковзним середнім наведено в табл. 6.

Аналіз отриманих результатів дозволяє зробити такі висновки:

Адекватна модель повинна обов'язково мати ковзне середнє. З обчислених варіантів моделей кращою є АРКС((1,2,3), (1)), оскільки для неї $DW = 2,1976$, $R^2 = 1$ (табл. 6). Таким чином, модель, що адекватно описує дисперсію, можемо записати як

$$\varepsilon_1^2(k) = 0,1738 + 0,1180\varepsilon_1^2(k-1) - 0,0495\varepsilon_1^2(k-2) + 0,0369\varepsilon_1^2(k-3) + \varepsilon_2(k).$$

Таблиця 6. Варіанти оцінювання моделей з ковзним середнім

	(1)	(1, 2)	(1, 2, 3)	(1, 2, 3, 9)	(1, 2, 3, 9, 10)
β_0	0,1652 (2,49E+15)	0,1777 (1,98E+14)	0,1738 (1,48E+15)	0,1606 (2,56E+14)	0,1350 (3,54E+15)
β_1	0,1666 (5,59E+14)	0,1199 (2,32E+14)	0,1180 (1,84E+14)	0,1230 (3,56E+14)	0,1164 (5,82E+14)
β_2		-0,0263 (-4,33E+13)	-0,0495 (-6,19E+14)	-0,0458 (-9,58E+13)	-0,0429 (-1,11E+15)
β_3			0,0369 (9,27E+14)	0,0507 (1,86E+14)	0,0504 (1,62E+15)
β_4				0,0397 (1,72E+14)	0,0403 (1,34E+15)
β_5					0,1218 (3,75E+15)
β_6	4,87E-15 (16,0641)	8,13E-14 (15,7368)	-8,31E-15 (-12,9103)	-5,35E-14 (-15,4494)	-2,78E-15 (-13,7294)
RSS	5,06E-29	3,63E-28	1,09E-29	7,65E-28	1,29E-29
DW	1,7425	1,6296	2,1976	1,8840	1,5241
R ²	1,0000	1,0000	1,0000	1,0000	1,0000

Отримані моделі далі будуть використані для побудови прогнозу.

4. ПРОГНОЗУВАННЯ ПРОЦЕСІВ ЗА ДОПОМОГОЮ РІЗНИЦЕВИХ РІВНЯНЬ

Приклад 4.1. Отримання функції прогнозування за допомогою розв'язку різницевого рівняння для процесу АРІКС(1,1,2)

$$y(k) = y(k-1) + a_0 + \varepsilon(k) + \beta_1 \varepsilon(k-1) + \beta_2 \varepsilon(k-2).$$

Введемо позначення: $e(k) = \varepsilon(k) + \beta_1 \varepsilon(k-1) + \beta_2 \varepsilon(k-2)$ і перепишемо задане рівняння $y(k) = a_0 + y(k-1) + e(k)$, розв'язок якого має вигляд

$$y(k) = a_0 k + y_0 + \sum_{i=1}^k e(i). \text{ Звідси знайдемо } y_0: y_0 = y(k) - a_0 k - \sum_{i=1}^k e(i).$$

$$\text{Розв'язок для моменту } k+s: y(k+s) = y_0 + a_0(k+s) + \sum_{i=1}^{k+s} e(i).$$

Підставимо в цей розв'язок отриманий вище вираз для y_0 .

$$y(k+s) = y(k) - a_0 k - \sum_{i=1}^k e(i) + a_0 k + a_0 s + \sum_{i=1}^{k+s} e(i) = y(k) + a_0 s + \sum_{i=1}^s e(k+i).$$

$$\text{Оскільки } \sum_{i=1}^s e(k+i) = \sum_{i=1}^s \varepsilon(k+i) + \beta_1 \sum_{i=1}^s \varepsilon(k+i-1) + \beta_2 \sum_{i=1}^s \varepsilon(k+i-2),$$

ТО

$$y(k+s) = y(k) + a_0 s + \sum_{i=1}^s \varepsilon(k+i) + \beta_1 \sum_{i=1}^s \varepsilon(k+i-1) + \beta_2 \sum_{i=1}^s \varepsilon(k+i-2).$$

На основі останнього рівняння запишемо функцію прогнозування на один крок: $E_k[y(k+1)] = a_0 + y(k) + \beta_1 \varepsilon(k) + \beta_2 \varepsilon(k-1)$, оскільки $E_k[\varepsilon(k+i)] = 0, \forall i > 0$.

Функція прогнозування на два кроки

$$E_k[y(k+2)] = 2a_0 + y(k) + (\beta_1 + \beta_2)\varepsilon(k) + \beta_2 \varepsilon(k-1),$$

і для довільного числа кроків прогнозування маємо

$$E_k[y(k+s)] = sa_0 + y(k) + (\beta_1 + \beta_2)\varepsilon(k) + \beta_2 \varepsilon(k-1).$$

Отримане рівняння — це рівняння прямої (її нахил визначається коефіцієнтом a_0), на яку накладається зважений випадковий процес. Для порівняння точності прогнозу на основі різницевих рівнянь скористаємось фільтром Калмана та методом, запропонованим в пакеті Eviews.

Критерії аналізу точності прогнозу. Існує багато методів, які можна застосувати для порівняння прогнозуючих моделей. Один з критеріїв, що використовується найчастіше, є середньоквадратична похибка або її квадратний корінь. Цей критерій представляє собою суму квадратичних відхилень прогнозованих значень від досліджуваних.

Корінь з середньоквадратичної похибки, що використовує N прогнозованих значень, має форму

$$\text{RMSE}(\hat{y}^2(k)) = \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{y}^2(k) - y^2(k))^2}.$$

Інший спосіб визначення якості прогнозуючої моделі базується на модулі відносного відхилення від істинного значення. Усереднений модуль відносної похибки обчислюється за формулою

$$\text{MAPE}(\hat{y}^2(k)) = \frac{1}{N} \sum_{k=1}^N \frac{|\hat{y}(k) - y(k)|}{|y(k)|}.$$

Міра стійкості до відхилень від усталеного стану виражена за допомогою медіани квадратичної похибки

$$\text{MedSE}(\hat{y}^2(k)) = \text{Median} \left(\hat{y}^2(k) - y^2(k) \right)^2.$$

Ці три критерії були порівняні між собою з використанням індексу працездатності, взятого з теорії прийняття рішень (правило Севіджа-Ньюханса)

$$\text{Perf}_i = \sum_{j=1}^n \frac{EC_i - \min_i(EC_i)}{\min_i(EC_i)},$$

де EC — один з критеріїв, описаних вище; n означає кількість акцій, для яких отримані прогнози. Він може бути інтерпретований як відносна втрата точності однією з моделей порівняно з моделлю, постфактум кращою для акції j .

Зручним методом контролю якості є використання регресії

$$y^2(k) = \nu + \omega \hat{y}^2(k) + \varepsilon(k).$$

Розглянемо конкретні приклади застосування розробленої методики для побудови прогнозу поведінки дисперсії ряду.

Приклад 4.2. Розглянемо ряд, що описує дисперсію вартості акцій однієї з провідних компаній на Нью-Йоркській фондовій біржі. На основі отриманої вище моделі, яка адекватно описує процес

$$\begin{aligned} \varepsilon_1^2(k) = & 0,0714 + 0,1187 \varepsilon_1^2(k-1) - 0,1123 \varepsilon_1^2(k-2) - \\ & - 0,096 \varepsilon_1^2(k-4) + 0,1257 \varepsilon_1^2(k-5) + \varepsilon_2(k), \end{aligned}$$

побудовано прогноз вартості акцій на 5 кроків та порівняно його з реальними значеннями ряду. Графік прогнозу наведено на рис. 4.

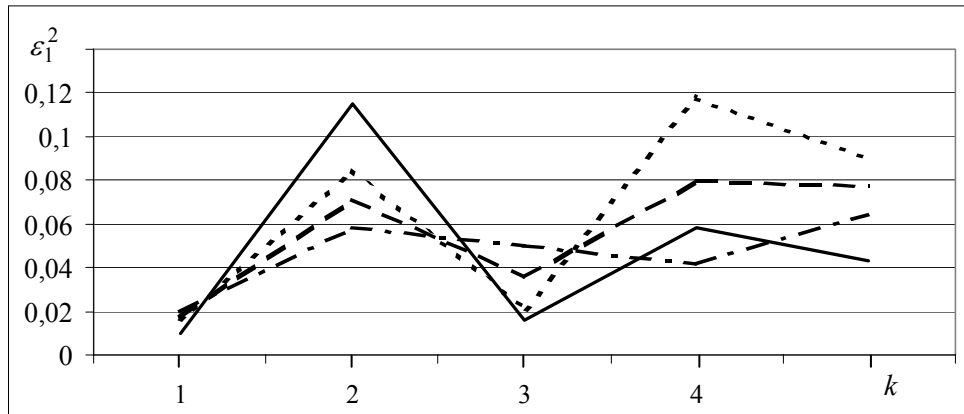


Рис. 4. Порівняльний графік прогнозу дисперсії вартості акцій однієї з провідних компаній на Нью-Йоркській фондовій біржі: — істинні значення; на основі моделі; - · - · - фільтр Калмана; - - - - метод подібних траєкторій

Статистичні параметри отриманого прогнозу наведені в табл. 7.

Таблиця 7. Оцінка якості прогнозу

Метод прогнозування	Максимальне відхилення		Мінімальне відхилення		Сума квадратів похибок
	Абсолютне	$\Delta \varepsilon_1^2, \%$	Абсолютне	$\Delta \varepsilon_1^2, \%$	
За моделлю	0,055	73	0,003	8	0,0338
Фільтр Калмана	0,059	78	0,007	12	0,0337
Метод подібних траєкторій	0,058	77	0,007	10	0,0335

Приклад 4.3. Побудуємо прогноз поведінки ряду, що відображає дисперсію вартості акцій УКРНАФТА. Адекватна модель, яка описує дисперсію, має вигляд

$$\varepsilon_1^2(k) = 0,1738 + 0,1180\varepsilon_1^2(k-1) - 0,0495\varepsilon_1^2(k-2) + 0,0369\varepsilon_1^2(k-3) + \varepsilon_2(k).$$

Знайдемо однорідний розв'язок: $\alpha^3 - 0,118\alpha^2 + 0,0495\alpha - 0,0369 = 0$.

Оскільки дане рівняння має один дійсний корінь ($a_1 = -0,4311$) та два комплексні, повний розв'язок має вигляд

$$\varepsilon_1^2(k) = y^{pd} + A\alpha^k + \beta_1 r^k \cos(k\theta + \beta_2) + y^{ps},$$

$$\text{де } r = \sqrt{a_2'}, \quad \cos(\theta) = \frac{a_1'}{2\sqrt{a_2'}}.$$

Значення a_1' та a_2' знайдемо так: з рівняння $\alpha^3 = a_1\alpha^2 - a_2\alpha - a_3 = 0$ знайдемо значення дійсного кореня α_1 і запишемо: $(\alpha - \alpha_1) \times (\alpha^2 + a_1'\alpha + a_2') = 0$. Розкривши дужки, виразимо a_1' , a_2' через значення a_1 і a_2 .

$$\alpha^3 + a_1'\alpha^2 + a_2'\alpha - \alpha^2\alpha_1 - \alpha a_1'\alpha_1 - a_2'\alpha_1 = 0,$$

$$\alpha^3 + (a_1' - \alpha_1)\alpha^2 + (a_2' - a_1'\alpha_1)\alpha - a_2'\alpha_1 = 0,$$

а звідси $a_1' = \alpha_1 + a_1 = -0,5491$, $a_2'a_1^2 + a_1a_1 + a_2 = 0,2862$.

Таким чином $r = 0,535$; $\theta = 2,11$.

Частковий розв'язок для детермінованої частини:

$$y^{pd} = 0,1738 - 0,1180y^{pd} + 0,0495y^{pd} - 0,0369y^{pd} \quad \text{або} \quad y^{pd} = 0,2634.$$

Частковий розв'язок для стохастичної частини:

$$y^{ps} = \sum_{i=0}^{\infty} \beta_i \varepsilon(k-i), \quad \text{де } \beta_i = a_1^i, \quad i = 0, 1, 2, \dots$$

Повний розв'язок: $\varepsilon_1^2(k) = y^{pd} + A\alpha^k + y^{ps}$.

Для того щоб знайти значення невідомої констант A , β_1 , β_2 скористаємось початковими умовами: $\varepsilon_1^2(0) = 4,809$; $\varepsilon_1^2(1) = 1,659$; $\varepsilon_1^2(2) = 0,450$. В результаті отримаємо $A = 6,6272$; $\beta_1 = -3,8771$; $\beta_2 = 1,0041$.

Отже повний розв'язок має вигляд

$$\varepsilon_1^2(k) = 0,2634 + 6,6272(-0,4311)^k - 3,8771 \cdot 0,535^k \cos(k \cdot 2,11 + 1,0041) + \sum_{i=0}^{\infty} (-0,8513)^i \varepsilon(k-i).$$

Отриманий розв'язок свідчить про присутність гармонійного коливального процесу, який відповідає реальним коливанням цін акцій компанії УКРНАФТА. Загалом розв'язок має збіжний характер, оскільки $\alpha = -0,4311$; $r = 0,535$, тобто обидва значення є меншими одиниці за модулем.

Таким чином, прогнозоване значення математичного сподівання дисперсії на s періодів дискретизації на основі отриманого розв'язку можна записати як

$$\begin{aligned} \varepsilon_1^2(k+s) + 0,2634 + [y(k) + 0,28]/0,1858(-0,4311)^k - \\ - [y(k-1) + 6,0227]/0,5348 \times \\ \times 0,535^s \cos[s \cdot 2,11 - \arccos(y(k-2) - 6,8906)/3,8771] + \\ + \sum_{i=0}^{s-1} (-0,8513)^i \varepsilon(k+s-i). \end{aligned}$$

Побудуємо прогноз вартості акцій на 5 кроків та порівняймо його з реальними значеннями ряду. Графік прогнозу наведено на рис 5.

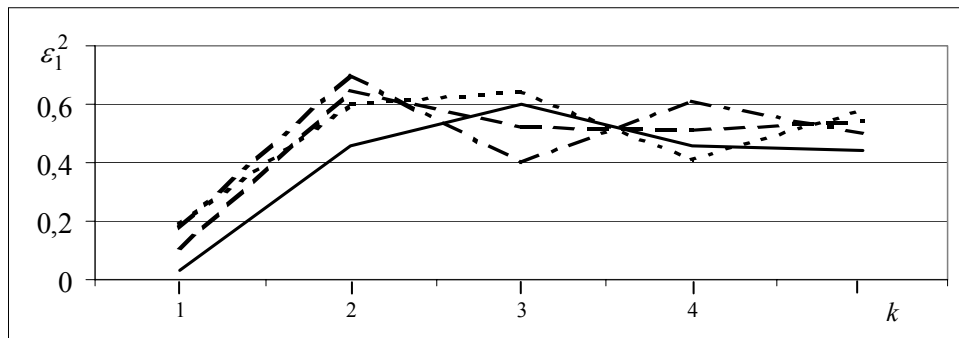


Рис. 5. Порівняльний графік прогнозу поведінки ряду, який описує дисперсію вартості акцій УКРНАФТА: — істинні значення; на основі прогнозуючої функції; - · - · - · повного розв'язку; - - - - метод подібних траєкторій

Статистичні параметри отриманого прогнозу наведені в табл. 8.

Таблиця 8. Характеристики якості прогнозу

Метод прогнозування	Максимальне відхилення		Мінімальне відхилення		Сума квадратів похибок
	Абсолютне	$\Delta \varepsilon_1^2, \%$	Абсолютне	$\Delta \varepsilon_1^2, \%$	
Прогнозуюча функція	0,2418	39,19	0,0465	6,95	0,1068
На основі розв'язку	0,2494	40,24	0,0316	9,47	0,1682
Метод подібних траєкторій	0,1486	36,45	0,0263	5,74	0,1174

Таким чином, запропонований системний підхід може бути використаний при побудові математичних моделей процесів за часовими рядами та прогнозів поведінки рядів, зокрема, гетероскедастичних та нелінійних щодо змінних процесів. Для більшості моделей кращим є прогноз на основі моделі, але при прогнозуванні на один крок кращим є прогноз, побудований за допомогою фільтра Калмана, що пояснюється вибором оптимального коефіцієнта фільтра.

ВИСНОВКИ

Розроблена методика ітеративної побудови моделей гетероскедастичних процесів на основі часових рядів, яка дає можливість отримати модель з високим ступенем адекватності для опису динаміки дисперсії часового ряду. Запропоновано метод подальшого підвищення адекватності моделей гетероскедастичних процесів у випадку присутності імпульсних адитивних складових часового ряду шляхом врахування цих складових при оцінюванні параметрів моделі. При цьому незначне ускладнення процедури оцінювання параметрів моделі дозволяє суттєво поліпшити її якість. Наведено приклади використання запропонованої методики опису гетероскедастичних процесів до прогнозування конкретних фінансово-економічних процесів. Отримані результати прогнозування свідчать про їх високу якість, що дозволяє скористатися ними при прийнятті рішень оперативного управління.

ЛІТЕРАТУРА

1. *Ивахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. — Киев: Наук. думка, 1982. — 296 с.
2. *Себер Дж.* Прикладной регрессионный анализ. — М.: Мир, 1982. — 450 с.
3. *Бард Й.* Нелинейное оценивание параметров. — М.: Статистика, 1979. — 349 с.
4. *Дрейпер Н., Смит Г.* Прикладной регрессионный анализ. — М.: Финансы и статистика, 1986. — 2. — 366 с.
5. *Бідюк П.І., Зворигіна Т.А.* Структурний аналіз методики побудови регресійних моделей за спостереженнями часового ряду // УСiМ. — 2003. — № 3. — С. 93–99.
6. *Бідюк П.І., Половцев О.В.* Аналіз та моделювання економічних процесів перехідного періоду. — Київ: НТУУ «КПІ», 1999. — 230 с.
7. *Бідюк П.І., Баклан І.В., Рифа В.Н.* Системний підхід к построению регрессионной модели по временным рядам // Системні дослідження та інформаційні технології. — 2002. — № 3. — С. 114–131.
8. *Enders W.* Applied econometric time series. — New York: John Wiley & Sons, Inc., 1995. — 434 p.
9. *Johnston J., DiNardo J.* Econometric methods. — New York: McGraw—Hill, Inc., 1997. — 530 p.

Надійшла 25.07.2003