

Моделювання і прогнозування гетероскедастичних процесів

Розглядається методика побудови математичних моделей гетероскедастичних процесів і її застосування до описання 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.

A methodology of model constructing for heteroscedastic processes is considered and its application to describing dynamics of time series. A simplified test for heteroscedasticity is proposed and algorithm for model improvement thanks to taking into consideration the spikes that are substantially greater than mean value of a series. The variance forecasting functions are constructed as a measure of risk on the basis of equations solutions. Some examples of forecasting real time series are provided.

Вступ

Гетероскедастичні процеси, тобто процеси із змінною в часі дисперсією, – поширений клас фінансово-економічних процесів, особливо в нестійкій перехідній економіці.

За визначенням такі процеси відносяться до класу слабо нестационарних процесів, а тому описання їх динаміки має деякі особливості у порівнянні із стационарними.

Одним із популярних підходів до моделювання нестационарних процесів є метод групового врахування аргументів [1], який був успішно застосований до побудови моделей ряду 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 спершу 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 оправданих обчислень. Найбільш зручним представляється тест, наведений нижче.

Спрощений тест на гетероскедастичність. Спрощений тест на гетероскедастичність складається з наступних кроків: 1) оцінити авторегресію $y(k) = a_0 + a_1 y(k-1) + \varepsilon(k)$ або більш високого порядку, наприклад другого чи третього; 2) побудувати ряд $\{\varepsilon^2(k)\}$, скориставшись залишками від оцінювання попередньої моделі; 3) оцінити регресію: $\varepsilon^2(k) = \alpha_0 + \alpha_1 [0,4 \varepsilon(k-1) + 0,3 \varepsilon(k-2) + 0,2 \varepsilon(k-3) + 0,1 \varepsilon(k-4)]$; 4) якщо коефіцієнт α_1 відмінний від нуля в статистичному смислі, тобто є значимим, то отримана модель для $\varepsilon^2(k)$ описує

гетероскедастичний процес.

Оскільки в цій моделі оцінюються тільки коефіцієнти α_1 і α_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 гіпотези, 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 АРУГ(p, q) на основі максимізації функції правдоподібності $\sum a_i y(k-i) + \varepsilon_1(k)$

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

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

та псевдолінійних моделей АРУГ

де $\varepsilon_1(k)$ – послідовність k -випадкових величин із змінною в часі дисперсією; $\varepsilon_2(k)$ – послідовність випадкових величин з наступними властивостями: $E[\varepsilon_2(k)] = 0$,

$$J = \sum_{k=1}^N [y(k) - \hat{y}(k)]^2$$

необхідно знайти:

1) оцінку вектора параметрів $\theta_1 = [a_0 \ a_1 \ \dots \ a_p]^T$, мінімізуючи критерій $\hat{y}(k) = \sum_{i=1}^p a_i y(k-i) + \varepsilon_1(k)$

де

$$L(\theta_2) = \log(f_{y_1}(y_{1m}; \theta_2)) \ ;$$

2) оцінку вектора $\theta_2 = [\alpha_0 \ \alpha_1 \ \beta_1 \ \dots \ \beta_m]^T$ із умови $\prod_{k=1}^N (f_{\varepsilon_1}(\varepsilon_1(k); \theta_2) \prod_{i=1}^m f_{\varepsilon_2}(\varepsilon_2(k-i); \theta_2))$ правдоподібності

$$\hat{\varepsilon}^2(k+s|k) = E_k[\varepsilon^2(k)], \quad s=1,2,\dots$$

3) на основі оціненої моделі знайти оцінку прогнозу дисперсії:

$$J_1 = E[\hat{\varepsilon}^2(k+s) \hat{\varepsilon}^2(k+s)] \rightarrow \min$$

використовуючи всю наявну інформацію на момент часу k , із умови

2. Методика побудови моделей гетероскедастичних процесів

Для прикладу розглянемо спочатку узагальнену авторегресійну модель гетероскедастичного процесу (УАРУГ). При використанні такої моделі умовна дисперсія процесу $\varepsilon(k)$ визначається $h(k)$ допомогою моделі авторегресії з ковзним середнім (АРКС). Нехай похибки моделі описуються рівнянням:

$$(1)$$

де $v(k)$ процес білого шуму з одиничною дисперсією $\sum_{i=1}^k \varepsilon^2(k) = \sum_{i=1}^k v^2(k)$ умовна дисперсія $i=1$ процесу, яка дозволяє побудувати узагальнену умовно гетероскедастичну модель (УАРУГ):

$$(2)$$

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

Математичне сподівання для $\varepsilon(k)$:

Важливим моментом є те, що умовна дисперсія процесу $\{\varepsilon(k)\}$ є залежною 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 скінченою, необхідно, щоб корені характеристичного рівняння, записаного для (2), знаходились всередині кола

одиночного радіуса на комплексній площині. Таким чином, основною відмінною властивістю моделі ОАРУГ є те, що умовна дисперсія збурень, які діють на процес $\{y(k)\}$, є процесом авторегресії з ковзним середнім.

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

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

- якщо не вдається побудувати адекватну модель АРКС, то, використовуючи автокореляційну функцію для квадратів залишків, необхідно побудувати модель УАРУГ, яка базується на проведенні аналізу поведінки дисперсії процесу; корелограма процесу $\{\varepsilon^2(k)\}$ дає можливість визначити наявність гетероскедастичності.

Оскільки $E_{k-1}[\varepsilon^2(k)] =$

$a_0 + \sum_{i=1}^q \alpha_i \varepsilon^2(k-i) + \sum_{i=1}^p \beta_i h(k-i)$
то рівняння (2) можна записати у формі:

(3)

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

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 дисперсії його необхідно видалити, наприклад, шляхом обчислення перших різниць або різниць вищого порядку. Досить часто візуальний аналіз даних дозволяє отримати суттєву інформацію щодо присутності гетероскедастичності. Разом з візуальним

аналізом необхідно розглядати параметри описової статистики, які полегшують визначення структури моделі.

Крок 2. Використовуючи АРКС та ЧАКФ для експериментальних даних, побудувати модель АР(p) або АРКС(p, q) 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 $\hat{\varepsilon}(k)$:

$$d_{\hat{\varepsilon}}(s) = \frac{\sum_{k=s+1}^N [\hat{\varepsilon}^2(k) - \hat{\sigma}_\varepsilon^2][\hat{\varepsilon}^2(k-s) - \hat{\sigma}_\varepsilon^2]}{\text{число залишків після побудови моделі АР чи АРКС} [\hat{\varepsilon}^2(k) - \hat{\sigma}_\varepsilon^2]}$$

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

(4)

Якщо існують такі значення $\rho(s)$, які відрізняються від нуля (в певному смислі), то це свідчить про присутність процесу АРУГ або УАРУГ. Для того, щоб переконатись у присутності гетероскедастичності, використовують Q -статистику Льюнга-Бокса, яка обчислюється за виразом [8]:

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

де $n = N/4-1$ (емпіричне значення). Якщо значення $\hat{\varepsilon}(k)$ некорельовані, то Q -статистика повинна мати розподіл χ^2 з n ступенями свободи.

Крок 4. Побудувати модель УАРУГ (або іншу модифікацію)

(5)

використовуючи ряд значень $\hat{\varepsilon}^2(k)$ і ряд значень умовної дисперсії $h(k)$. Якщо в цій моделі хоча б один із коефіцієнтів α_i , $i=1, \dots, q$ є значимим, то процес є дійсно гетероскедастичним. Оскільки модель (3) 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 $AR(p)$ чи $AR(p, q)$, 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 $AR(p)$ чи $AR(p, q)$. 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 описання даних, що мають надлишковий куртозис. Стандартну модель УАРУГ можна використати також для описання похибок моделювання, що мають t -розподіл (модель УАРУГ- t). Однак існують ще й інші проблеми з досягненням високого ступеня адекватності моделей такого типу, наприклад, у тих випадках, коли ряд містить адитивні складові (імпульси), які суттєво перевищують поточне середнє значення, а також ближні сусідні значення. Така проблема досить часто виникає при побудові моделей цін акцій на біржах, а також при моделюванні доходів від акцій. Тому в даному параграфі пропонується 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 авторегресії з ковзним середнім $AR(p, q)$, тобто

$$y(k) = a_0 + a_1 y(k-1) + \dots + a_p y(k-p) + b_1 \varepsilon(k) + b_2 \varepsilon(k-1) + \dots + b_q \varepsilon(k-q), \quad (6)$$

де $\varepsilon(k)$ – випадкова послідовність типу білого шуму з нульовим середнім та скінченною дисперсією σ_ε^2 . В моделі (6) відсутня константа a_0 , оскільки припускається, що вона будується на основі відхилень 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 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, що враховує значні адитивні імпульсні складові часового ряду, має вигляд:

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

де

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

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

При оцінюванні моделі $AR(p, q)$ на основі такого ряду залишки моделі можна описати наступним чином:

$$\hat{\varepsilon}(k) = \begin{cases} 0, & k < l, \\ \omega(k), & k = l, \\ -D(j)\omega(k-l-j), & k > l, \end{cases} \quad (8)$$

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

$$\omega(k) = \begin{cases} 0, & k < l, \\ 1, & k = l, \\ -D(j)\omega(k-l-j), & k > l, \end{cases} \quad (9)$$

яке можна розглядати як модель регресії для $\hat{\varepsilon}(k)$:

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

де

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

і

Тепер вплив суттєвого імпульсного значення на модель в момент $k = l$ можна представити у вигляді:

(11)

Функцію $\hat{\beta}(l)$ можна представити 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 ряду 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 через $\hat{\sigma}_e^2$, то можна записати стандартний статистичний параметр (статистику) у вигляді:

(12)

Вплив імпульсного значення ряду вважається значним, якщо $\hat{t} > c$, де c – деяке порогове значення. Оскільки в асимптотиці \hat{t} має нормальний розподіл, то порогове значення c можна визначити з таблиць для t -розподілу. На основі емпіричних досліджень пропонується використовувати для цього величину $c = 4$ [9]. 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

Розглянутий підхід до врахування імпульсних значень часового ряду даних поширюється на логарифмізовані дані множини імпульсів. Тобто процедура повторюється для кожного імпульсу до тих пір, 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

і запишемо можливу структуру узагальненої авторегресійної умовно гетероскедастичної моделі у вигляді:

(13)

(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) являють собою модель УАРМГ. Але якщо $\eta(k)$ має умовний t -розподіл, то модель (13), (14) називають УАРУГ- t . Для покращення описання ряду з імпульсними значеннями скористаємося стандартною моделлю УАРУГ, але будемо коригувати її із врахуванням наявності імпульсних значень в ряді $\eta(k)$.

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

(15)

де $x(k) \equiv d^2(k) - h(k)$. Рівняння (15) – це модель $(L) \equiv (1 - \hat{A}_1^* z^{-1}) / (1 - \hat{A}_1 z^{-1}) d^2(k)$. Зазначимо, що ряд $x(k)$ можна розглядати як гетерогенний ряд залишків. Тепер алгоритм побудови моделі із врахування наявності імпульсних значень можна представити у вигляді наступних кроків:

Крок 1. Обчислити оцінки параметрів рівнянь (13), (14) методом максимальної правдоподібності. В результаті отримаємо оцінки $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\gamma}_0$, 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 наявності імпульсних значень.

Крок 2. З рівняння (8) випливає, що

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 при $k = l$. Для того щоб врахувати змінну в часі дисперсію величини $\hat{x}(k)$, її необхідно обчислювати для вибраного рухомого вікна, а не для вибірки в цілому. Статистики \hat{t} , що визначаються виразом (12), обчислюються для $\hat{x}(k)$ і порівнюються із значенням $c = 4$.

Крок 3. Якщо обчислення $\hat{c}(k)$ при $k = l$ та найбільшому значенні статистики \hat{t} , яке повинне перевищувати $c = 4$, замінюється на $\hat{x}^*(k)$ (оскільки це показано в рівнянні (7)) з використанням 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:

для моменту часу $k = l$. Скоригувати ряд даних відносно імпульсних значень наступним чином:

(16)

Наведений вираз показує, що

при $k = l$. З цього випливає, що для членів ряду, які коригуються у відповідності до наявності імпульсного значення, знак не змінюється.

Крок 5. Після обчислення значень ряду $d^*(k)$ перейти на крок 1 і повторювати всю наведену процедуру до тих пір, поки $\hat{t} > 4$, тобто до тих пір, поки не будуть оброблені всі члени ряду, що 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 $\hat{h}^*(N+1)$, за допомогою рівняння (14).

Якщо $d(k)$ в рівнянні (13) представляє собою корельований процес, то перед використанням запропонованого п'ятикрокового алгоритму необхідно фільтрувати детерміновану динаміку процесу за допомогою моделі АР або АРКС (тобто спочатку побудувати модель такого типу). 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 b 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 доход російської біржі сільгосппродуктів за 1996-1999 роки. Ряд $d(k)$ представлений на рис. 1. Статистичні дані включає оцінювання параметрів математичних моделей на

основі щотижневих $\eta(k)$ членів $M(0,1)$ річних даних, які використовуються для $h(k)$ та оцінки $\varepsilon_1^2(k)$ в однокроковому 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 коригування впливу імпульсних значень застосовується послідовно до всіх значень вибірки.

На основі даного часового ряду оцінено моделі $AR(p)$, потім $УАРУГ(1,1)$, а також модель для доходу, що має структуру:

Таблиця 1

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

Вибірка даних	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

Таблиця 2

Результати оцінювання моделей $AR(p)$ та $УАРУГ(1,1)$ після корекції

Вибірка даних	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 таблицею 1 оцінки параметрів α_1 і γ_1 змінились суттєво. Наприклад, для періоду даних 1996-1999 значення α_1 зменшилось від 0,274 до 0,118, а значення γ_1 збільшилось від 0,468 до 0,694. Однак сума коефіцієнтів $\alpha_1 + \gamma_1$ суттєво не змінилась, 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. Процес функціонування виробничої фірми схематично зображено на рис. 1.

Нехай $q(k)$ – прибуток виробничої фірми 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 умовна дисперсія ціни на продукцію фірми в момент k , яка визначається на основі інформації на 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 доходу фірми.

Для описання прибутку вибрана наступна структура моделі:

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

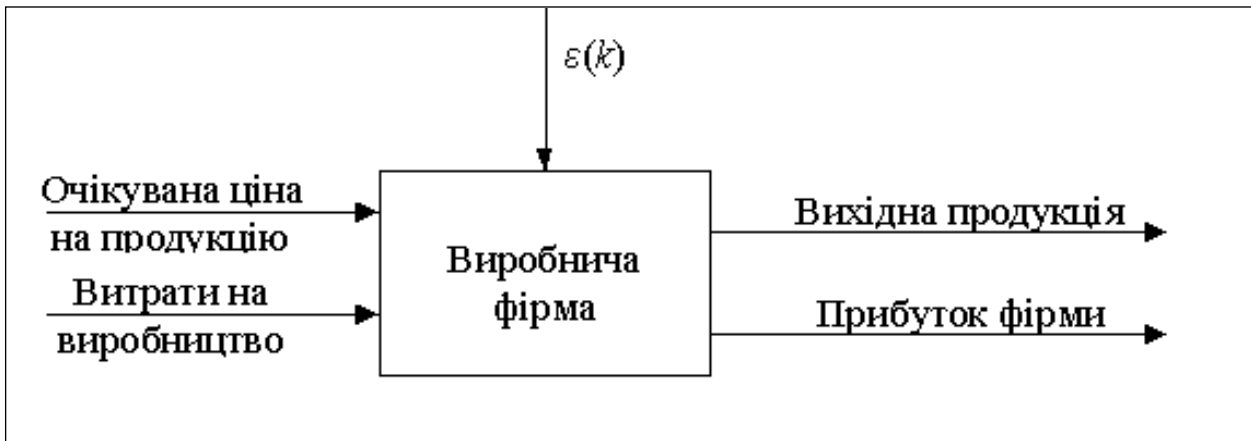


Рис. 1. Вхідні та вихідні змінні виробничої фірми

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 – вплив умовної дисперсії ціни продукції на прибуток фірми. Змінна $p^e(k)$ – оцінюється на основі ціни в попередньому кварталі.

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

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

Передбачається, що дисперсія ціни – гетероскедастичний процес:

$$(1 - 0,511L - 0,129L^2 - 0,130L^3 - 0,138L^4) p(k) = 1,632 + \varepsilon_2(k). \quad (19)$$

Тепер покладемо $p^e(k) = p(k)$, щоб підставити в (18), і в результаті отримаємо модель

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

Значення, обчислені в (19) і (20) підставляються в (17) для оцінки рівня ціни на продукцію. Коефіцієнти в (19) і (20) відносяться до прибутку фірми по виробництву – бройдерів. Рівняння для обчислення поточного прибутку:

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

Приклад 3.3. Ціни акцій на біржі. Розглянемо ряд $\varepsilon^2(k)$ для моделі, побудованої в попередньому розділі (приклад 1) АР (АР(1,5,8)) (коливання ціни акцій на біржі, рис. 2; часткова АКФ на рис. 3). В результаті дослідження АКФ і ЧАКФ встановлено наступне:

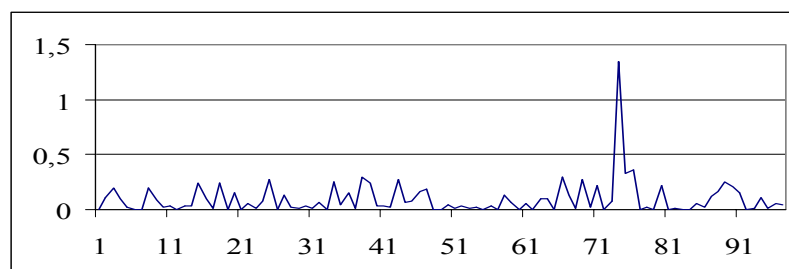


Рис. 2. Коливання ціни акцій на біржі

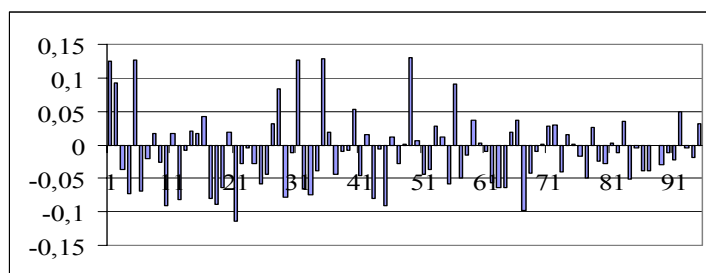


Рис. 3. Часткова автокореляційна функція ряду цін на біржі

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

3. Для створення математичного описання процесу необхідно розглянути

моделі $AR(1)$, $AR(2)$, $AR(3)$, $AR(4)$. Найбільш вірогідні номери запізнювань, що входять до складу моделі: 1, 2, 4, 5, 10. У табл. 3 наведено варіанти оцінювання декількох можливих структур регресійної

Таблиця 3

Варіанти оцінювання дисперсії вартості акцій однієї з компаній, що входили до числа провідних на Нью-Йоркській фондовій біржі за 1996 рік

	(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
R²	0,0155	0,0242	0,0307	0,0474	0,0536

З таблиці 3 випливає, що найкраща модель, яка описує дисперсію з обчислених варіантів, є $AR(1, 2)$, оскільки для неї $DW = 1,9899$, $R^2 = 2,4041$, $AIC = -0,7846$. 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 $R^2 = 0,0242$ показує, що модель можна покращити за

рахунок введення ковзного середнього. Варіанти оцінювання моделей з ковзним середнім наведені в табл. 4.

З таблиці 4 випливає, що краща модель повинна обов'язково включати в себе ковзне середнє. З обчислених варіантів моделей кращою є $ARPC((1, 2, 4, 5), (1))$, оскільки для неї $DW = 2,0675$, $R^2 = 1$, $AIC = -\infty$. Крім того,

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_4	-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
<i>DW</i>	0,4843	1,1944	1,2519	2,0675	2,0796
R^2	0,9970	0,9980	0,9980	0,9990	0,9990

велике значення t -статистики для оцінок коефіцієнтів свідчить про їх значимість у статистичному сенсі. Таким чином, модель, що описує дисперсію можемо записати як

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

Приклад 3.4. Ціни акцій компанії УКРНАФТА. Розглянемо ряд $\varepsilon^2(k)$ для побудованої в попередньому розділі (приклад 2) моделі $AR(1,2,3)$ щоденної вартості акцій компанії УКРНАФТА за 2001 рік. В результаті дослідження АКФ і ЧАКФ встановлено наступне:

1. Значення коефіцієнтів АКФ такі: $\rho_1 = 0,166$; $\rho_2 = -0,006$; $\rho_3 = 0,028$; $\rho_9 = 0,038$; $\rho_{10} = 0,132$.

2. Значення коефіцієнтів ЧАКФ мали такі значення: $\Phi_{1,1} = 0,166$; $\Phi_{2,2} = -0,035$; $\Phi_{3,3} = 0,036$; $\Phi_{9,9} = 0,043$; $\Phi_{10,10} = 0,122$.

3. Для створення математичного описання процесу необхідно розглянути моделі $AR(1)$, $AR(2)$, $AR(3)$, $AR(4)$.

Найбільш вірогідні номери запізнювань, що входять до складу моделі: 1, 2, 3, 9, 10. У табл. 5 наведені варіанти оцінювання декількох можливих структур регресійної моделі.

Аналіз отриманих результатів дозволяє зробити такі висновки:

З таблиці 5 випливає, що найкраща модель, що описує дисперсію з обчислених варіантів, є $AR(1, 2)$, оскільки для неї $DW = 2,0029$, $R^2 = 0,0144$, $AIC = 1,4235$. Таким чином, модель, що описує процес, можемо записати як

$$\begin{aligned} \varepsilon_1^2(k) = & 0,1777 + 0,1199 \varepsilon_1^2(k-1) - \\ & - 0,0263 \varepsilon_1^2(k-2). \end{aligned}$$

Значення $RSS = 57,8552$ та низькі значення R^2 показують, що модель можна покращити за рахунок введення ковзного середнього. Оцінки параметрів моделей з ковзним середнім наведено в табл. 6.

Аналіз отриманих результатів дозволяє зробити такі висновки:

З таблиці 6 випливає, що адекватна модель повинна обов'язково включати в себе ковзне середнє. З обчислених варіантів

Таблиця 5

**Варіанти оцінювання дисперсії щоденної вартості акцій
компанії УКРНАФТА за 2001 рік**

	(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
<i>AIC</i>	1,3820	1,3818	1,3909	1,4151	1,4037
<i>BSC</i>	1,4098	1,4235	1,4468	1,4862	1,4892
<i>DW</i>	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

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
<i>DW</i>	1,7425	1,6296	2,1976	1,8840	1,5241
R^2	0,9960	0,9970	0,9970	0,9990	0,9990

моделей кращою є АРКС ((1,2,3), (1)), оскільки для неї $DW = 2,1976$, $R^2 = 1$, $AIC = -\infty$. Таким чином, модель, що описує дисперсію, можемо записати у вигляді

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

Отримані моделі далі будуть використані для оцінювання прогнозів.

4. Прогнозування процесів за допомогою різницевого рівняння

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

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

Введемо наступне позначення: $e(k) = \varepsilon(k) + \beta_1 \varepsilon(k-1) + \beta_2 \varepsilon(k-2)$ і перепишемо задане рівняння як

$$y(k) = a_0 k + y_0 + \sum_{i=1}^k e(i)$$

розв'язок якого має вигляд:

$$y_0 = y(k) - a_0 k - \sum_{i=1}^k e(i)$$

Знайдемо звідси y_0 :

$$y(k+s) = y_0 + a_0(k+s) + \sum_{i=1}^{k+s} e(i)$$

Розв'язок для моменту $k+s$:

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

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

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

Оскільки

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

то

На одній (останній) з рівнянь запишемо функцію прогнозування на один крок:

$$+ \beta_1 \varepsilon(k) + \beta_2 \varepsilon(k-1)$$

$$E_k[\varepsilon(k+i)] = 0, \quad \forall i > 0$$

оскільки

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

Функція прогнозування на два кроки:

$$(\beta_1 + \beta_2)\varepsilon(k) + \beta_2 \varepsilon(k-1)$$

і для $E_k[y(k+n)] = sa_0 + y(k) + n$ кроків прогнозування маємо:

$$+ (\beta_1 + \beta_2)\varepsilon(k) + \beta_2 \varepsilon(k-1)$$

Отримане рівняння – це рівняння прямої (її нахил визначається коефіцієнтом a_0), на яку накладається зважений випадковий процес. Для порівняння точності прогнозу на основі різницевого рівняння скористаємось фільтром Калмана та методом, запропонованим в системі 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 являє собою суму квадратичних відхилень прогнозованих значень від

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

Інший спосіб визначення якості прогнозуючої моделі базується на модулі відносного відхилення від істинного значення. Усереднений модуль відносної похибки обчислюється за наступною формулою [10]:

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

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

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

Зручним методом контролю якості прогнозу є регресування досліджуваних дисперсій за

прогнозованими, тобто використання регресії:

$$y^2(k) = \upsilon + \omega \cdot \hat{y}^2(k) + \varepsilon(k)$$

Для якісного прогнозу необхідно, щоб параметри цієї моделі мали такі значення: $\upsilon \rightarrow 0$, $\hat{\omega} \rightarrow 1$.

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

Приклад 4.2. Розглянемо ряд, що описує дисперсію вартості акцій однієї з компаній, що входить до числа провідних на Нью-Йоркській фондовій біржі. На основі моделі, що описує процес,

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

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

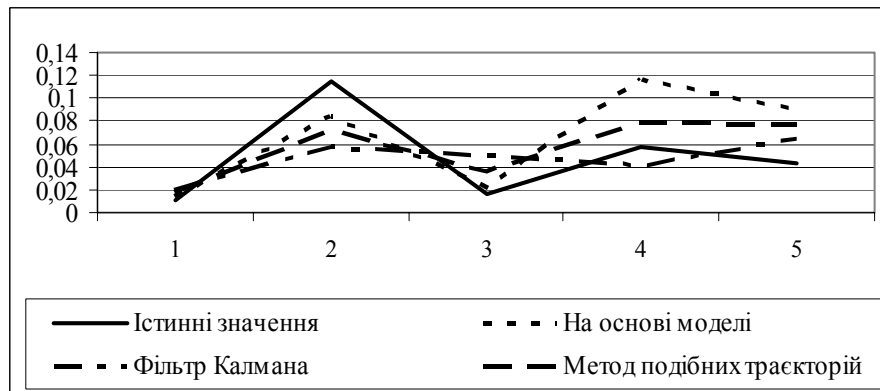


Рис. 4. Порівняльний графік прогнозу дисперсії вартості акцій однієї з компаній, що входить до числа провідних на Нью-Йоркській фондовій біржі

Таблиця 7

Метод прогнозування	Максимальне відхилення		Мінімальне відхилення		Сума квадратів похибок
	Абсолютне	%	Абсолютне	%	
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. Прогнозування поведінки ряду, що описує дисперсію вартості акцій УКРНАФТА.

Адекватна модель, що описує дисперсію, має наступний вигляд:

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

а) Знайдемо однорідний розв'язок характеристичного рівняння:

$$\alpha^3 - 0,118 \alpha^2 + 0,0495 \alpha - 0,0369 = 0.$$

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

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

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

Значення a_1' та a_2' знаходяться наступним чином: рівняння $\alpha^3 - a_1 \alpha^2 - a_2 \alpha - a_3 = 0$, знайшовши значення дійсного кореня α_1 , можемо записати: $(\alpha - \alpha_1)(\alpha^2 + a_1' \alpha + a_2') = 0$. Розкривши дужки, виразимо a_1' , a_2' через значення α_1 і a_2 :

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

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

а звідси

$$a_1' = \alpha_1 + a_1 = -0,5491,$$

$$a_2' = a_1^2 + a_1 a_2 + a_3 = 0,2862.$$

Звідси $r = 0,535$, $\theta = 2,11$.

б) Частковий розв'язок для детермінованої частини:

$$y^{pd} = 0,1738 - 0,1180 y^{pd} + 0,0495 y^{pd} - 0,0369 y^{pd},$$

$$\text{або } y^{pd} = 0,2634.$$

в) Частковий розв'язок для стохастичної частини:

$$y^{ps} = \sum_{i=0}^{\infty} \beta_i \varepsilon(k-i),$$

де $\beta_i = a_1^i$ $i = 0, 1, 2, \dots$

г) Повний розв'язок: $\varepsilon_1^2(k) = y^{pd} + A \alpha^k + y^{ps}$. Для того, щоб знайти значення невідомих 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 наступний вигляд:

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

Отриманий розв'язок свідчить про наявність гармонічного коливального процесу, що 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

$$\varepsilon_1^2(k+s) = 0,2634 + [y(k) + 0,28] / 0,1858 \times (-0,4311)^k - [y(k-1) + 6,0227] / 0,5348;$$

$$0,535^s \cos[s \times 2,11 + \arccos(y(k-2) - 6,8906) / 3,8771] + \sum_{i=0}^{\infty} (-0,8513)^i \varepsilon(k+s-i).$$

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

Статистичні параметри отриманого

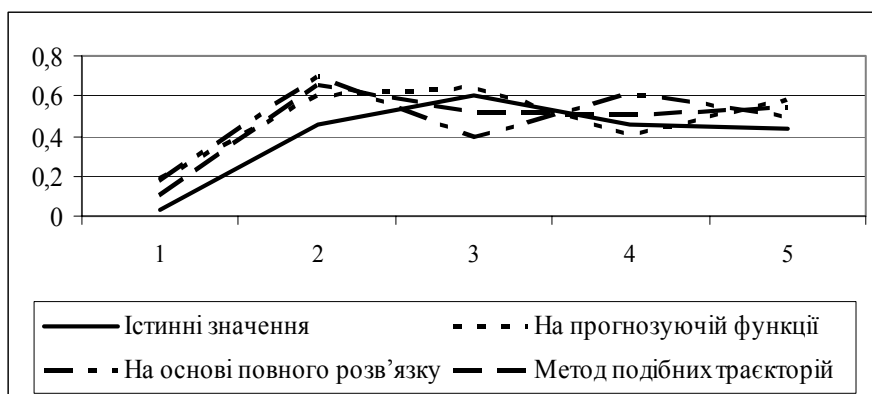


Рис. 5. Порівняльний графік прогнозу поведінки ряду, що описує дисперсію вартості акцій УКРНАФТА

Таблиця 8

Метод прогнозування	Максимальне відхилення		Мінімальне відхилення		Сума квадратів похибок
	Абсолютне	%	Абсолютне	%	
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 підхід може бути використаний при побудові математичних моделей процесів за часовими рядами та побудові прогнозів поведінки рядів, включаючи гетероскедастичні та нелінійні щодо змінних процеси. 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, ймовірно, використанням оптимального коефіцієнта фільтра та можливості врахування в алгоритмі фільтрації збурень стану і похибок статистичних даних.

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. Встановлено, незначне ускладнення процедури оцінювання параметрів моделі дозволяє суттєво підвищити її якість. Наведено приклади використання запропонованої методики описання гетероскедастичних процесів до прогнозування конкретних фінансово-економічних процесів. Отримані результати прогнозування свідчать про їх високу якість,

Література

1. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наукова 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 (т. 2). – М.: Финансы и статистика, 1986. – 366 с.
5. Бідюк П.І., Зворигіна Т.А. Структурний аналіз методики побудови регресійних моделей за спостереженнями часового ряду // Управляющие системы и машины. – 2003. – № 3.
6. Бідюк П.І. Системний підхід до прогнозування на основі моделей часових рядів // Системні дослідження та інформаційні технології. – 2003. – № 3. – С. 88-110.
7. Бідюк П.І., Половцев О.В. Аналіз та моделювання економічних процесів перехідного періоду. – Київ: НТУ КПІ, 1999. – 230 с.
8. Бідюк П.І., Рифа В.М. Системний підхід до побудови математичних моделей на основі часових рядів // Системні дослідження та інформаційні технології. – 2002. – № 3. – С. 114-131.
9. Enders W. Applied econometric time series. – New York: John Wiley & Sons, Inc., 1995. – 434 p.
10. Johnston J., DiNardo J. Econometric methods. – New York: McGraw-Hill, Inc., 1997. – 530 с.

Стаття надійшла до редколегії 25.12.2003 р.