ІНФОРМАЦІЙНЕ ТА МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ ЕКОНОМІЧНИХ ПРОЦЕСІВ.

INFORMATION AND MATHEMATICAL SUPPORT OF ECONOMIC PROCESSES

УДК 330.42

ДОСЛІДЖЕННЯ ІНДЕКСУ ПФТС ФОНДОВОГО РИНКУ УКРАЇНИ

В.М. Андрієнко, к.е.н., доцент

О.Г. Спиваков, магістр

Одеський національний політехнічний університет, Одеса, Україна

Андрієнко В.М., Співаков О.Г. Дослідження індексу ПФТС фондового ринку України.

Робота присвячена аналізу і прогнозуванню найважливішого макроекономічного показника України – індексу фондового ринку ПФТС. Застосований підхід до ідентифікації динаміки індексу з погляду присутності в їйнаслідківобурень, викликаних деякою зовнішньою шоковою дією.

Ключові слова: часовий ряд, стохастичний тренд, нейромережевий аналіз.

Андриенко В.М., Спиваков О.Г. Исследование индекса ПФТС фондового рынка Украины.

Работа посвящена анализу и прогнозированию важнейшего макроэкономического показателя Украины – индекса фондового рынка ПФТС. Применён подход к идентификации динамики индекса с точки зрения наличия в нём последствий возмущения, вызванных некоторым внешним шоковым воздействием.

Ключові слова: временной ряд, стахостический тренд, нейросетевой анализ.

Andrienko V.M., Spivakov O.G. Investigation index PFTS of fund market of Ukraine.

Work is devoted an analysis and prognostication of major macroeconomic index of Ukraine – to the index of fund market of PFTS. Approach is applied to authentication of dynamics of index from point of duration of being in her consequences of indignations, caused some external shock influence.

Keywords: temporal row, stochastic trend, analysis of neuron network.

Стан фондового ринку грає важливу роль для стабільного розвитку економіки. Крах фондового ринку, тобто сильне падіння (шок) курсової вартості цінних паперів за короткий проміжок часу, може викликати спад і депресію в економіці. Ухвалюючи рішення про інвестиції, фінансовий менеджер постійно оцінює поведінку в майбутньому, як окремих фінансових активів, так і ринку в цілому. Учасникові ринку потрібно хоч би приблизно представляти картину майбутнєго. Саме тому на перший план висувається завдання оцінки стану і тенденції розвитку ситуацій на фондовому ринку. З цією метою вся поточна і минула фінансова інформація ретельно аналізується за допомогою методів фінансового аналізу. Це дає розуміння минулого і поточного стану ринку. Проте яким би детальним не було це розуміння, прогнози, складені тільки на такого роду аналізі, не можуть служити надійною основою для ухвалення рішень про інвестування. В зв'язку з цим, побудова математичних моделей, що дозволяють краще зрозуміти структуру і поведінку ринку як єдиного цілого, так і його складових, довгий час привертали і продовжують привертати увагу дослідників і практиків. Проблема моделювання динаміки цін ринкових активів і їх прогнозування є достатньо складною, і її не можна назвати в даний час вирішеною.

Індекс українського фондового ринку ПФТС. Основними індикаторами фондового ринку є індекси, що розраховуються на підставі котирувань певної групи цінних паперів. Залежно від того, які компоненти містить індекс, він може відображати поведінку певної групи акцій або всього фондового ринку. Індекс Українського фондового ринку ПФТС (перша фондова торгова система) визнаний міжнародною фінансовою корпорацією (МФК) як єдиний індекс, використовуваний цією організацією при моніторингу внутрішнього стану українського фондового ринку. Індекс визначає середній рівень найбільш ліквідних українських акцій, які мають найбільшу ринкову капіталізацію. Інформацію про стан індексу ПФТС публікує на своєму сайті «Перша фондова торгова система», яка є електронною біржею цінних паперів України і підтримує роботу національної електронної системи торгівлі цінними паперами в режимі «online».

Математико-статистичні методи моделювання тимчасових рядів. Більшість математико-статистичних методів мають справу з моделями, в яких спостереження передбачаються незалежними і однаково розподіленими. При цьому основна увага приділяється проблемам ідентифікації моделей, відбору ендогенних і екзогенних показників, але майже не звертається уваги на формальний аналіз структури початкових статистичних рядів. Залежність між спостереженнями найчастіше розглядається як перешкода в ефективному застосуванні цих методів. Проте різноманітні дані в економіці, соціології, фінансах, комерції і інших сферах людської діяльності поступають у формі *тимчасових рядів*, в яких спостереження взаємно залежні, і характер цієї залежності якраз і представляє головний інтерес для дослідника. *Часовим рядом* називається сукупність спостережень економічного показника в різні моменти часу. Зазвичай часовий ряд розглядають як вибірку з послідовності випадкових величин X_t де t приймає цілочисельні значення від *1* до *T*. Сукупність випадкових величин $\{X_t, t \in [1,T]\}$ називають *дискретним випадковим процесом* або *стохастичним процесом*. Принципові відмінності тимчасового ряду від послідовності спостережень, утворюючих випадкову вибірку, полягають в наступному:

— на відміну від елементів випадкової вибірки члени тимчасового ряду не є незалежними;
— члени тимчасового ряду не обов'язково є однаково розподіленими.

Це означає, що властивості і правила статистичного аналізу випадкової вибірки не можна поширювати на тимчасові ряди. З іншого боку, взаємозалежність членів тимчасового ряду створює свою специфічну базу для побудови прогнозних значень аналізованого показника по спостереженням.

У 1938 році Вольд довів наступний фундаментальний результат [1]. Чисто недетермінований стаціонарний в широкому сенсі [2] випадковий процес X_t може бути представлений в наступному вигляді:

$$X_t - \overline{X}_t = \sum_{\tau=0}^{\infty} \psi_\tau \varepsilon_{t-\tau},$$

де X_t - математичне очікування цього процесу, а ε_t - білий шум з кінцевими математичним очікуванням і дисперсією. Тобто всякий стаціонарний в широкому сенсі випадковий процес представляється у вигляді лінійної комбінації білих шумів. При цьому повинна виконуватися умова $\sum_{t=0}^{\infty} |\psi_t| < \infty$. Оскільки реалізації білого шуму не спостережувані, вагові коефіцієнти визначені з точністю до множника. Без втрати спільності можна вважати, що $\psi_0 = 1$. Чим більше ваговий коефіцієнт ψ_t тим більше вплив випадкового обурень в момент $t-\tau$ на даний момент t. Виявилось, що у багатьох випадках досить розглядати не загальне представлення Вольда, а його окремі випадки, коли число доданків кінцеве. Такими окремими випадками є популярні в економетриці авторегресійні моделі *AR(p), MA(q), ARMA(p,q).*

Економічні показники, зокрема показники фондових ринків, не завжди поводяться стаціонарним чином. З макроекономіки відома їх *сезонна* і *циклічна* поведінка, крім того, вони можуть мати *тренд*. Часто ці види компонент присутні у ряді одночасного.

Аналітично часовий ряд можна виразити рівнянням вигляду:

$$X(t) = f(t) + S(t) + \varepsilon(t)$$

де $f(t)$ - тренд (довготривала тенденція) розвитку;
$S(t)$ - сезонна (періодична) компоненту;
$\varepsilon(t)$ - випадковий компонент.

Функція $f(t)$ визначає загальну тенденцію розвитку явища, що вивчається.

Для рядів із стохастичним трендом характерний необмежена спектральна щільність на низьких частотах поволі убуваюча автокореляційна функція. Такі ряди називають «тимчасові ряди з довготривалою кореляційною залежністю (time series with long memory)». У роботах зарубіжних учених, в першу чергу, C.W. Granger, J.R. Hosking, P.M. Robinson, R. Beran [4], був запропонований новий клас моделей *ARFIMA(p,d,q),* що допускає можливість нецілого параметра d і що отримав назву авторегресійний дріб, - інтегрований процес ковзаючого середнього. Характеристики таких тимчасових рядів володіють важливими властивостями, наприклад X_t є стаціонарним і оборотним при $d \in (-1/2, 1/2)$. При цьому X_t можна представити у вигляді [5]:

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t ,$$

де $B = X_{t-1} / X_t$ - оператор зрушення назад $\Phi\left(B\right) = 1 - \sum_{j=1}^{p} \phi_j B^j ; \Theta\left(B\right) = 1 + \sum_{j=1}^{q} \theta_j B^j$ - поліноми, передбачається, що все коріння рівнянь $\Phi\left(z\right) = 0,\ \Theta\left(z\right) = 0$ по модулю більше одиниці

$$(1-B)^d = 1 - dB + \frac{d(d-1)}{2} B^2 - ... = \sum_{j=0}^{\infty} \psi_j B^j ,$$

$$\psi_j = \prod_{0<k<j} \frac{k-1-d}{k} = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)},\ j = 0,1,2,... ,$$

де ε_t - гаусівський білий шум $WN(0,\sigma^2)$;

$\Gamma(x)$ – гамма-функція, яка визначається формулою

$$\Gamma(x) = \begin{cases} \int_0^{\infty} t^{x-1} e^{-t} dt, & x > 0, \\ x^{-1}\Gamma(1+x), & x < 0, x \neq -1,-2,... \end{cases} .$$

У окремому випадку, при $p,q = 0$ і $d \in (-1/2, 1/2)$ процес має вигляд:

$$X_t = (1-B)^{-d} \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} ,$$

де $\psi_j = \dfrac{\Gamma\left(j+d\right)}{\Gamma\left(j+1\right)\Gamma\left(d\right)}$.

Оцінку параметра d можна отримати з рівності $d = H - 0.5$ де H - показник Херста [6].

Значення $0.5 \le H \le 1$ має на увазі персистентний часовий ряд, а такий ряд характеризується ефектами довготривалої пам'яті. Теоретично це означає – що відбувається сьогодні, впливає на майбутнє. Така довготривала пам'ять має місце незалежно від масштабу часу: всі щоденні зміни співвідносяться зі всіма майбутніми щоденними змінами, всі щотижневі зміни співвідносяться зі всіма майбутніми щотижневими змінами. Персистентний часовий ряд є найпоширенішим типом, що зустрічається в природі. Він також є найпоширенішим і в економіці, і на фондових ринках. Такий ряд називають дробовий броунівський рух або узагальнений броунівський рух.

Значення $0 \le H \le 0.5$ означає антіперсистентність. Антіперсистентная система проходить меншу відстань, ніж випадкова система. Щоб система пройшла меншу відстань, вона повинна мінятися частіше, ніж імовірнісний процес.

Значення H=0.5 відповідає звичайному білому гаусіовському шуму $\left\{\varepsilon_t, t \ge 0 : \varepsilon_t \sim iid\ N(\mu,\sigma)\right\}$ випадковому блуканню броунівської частинки, тобто процесу повністю позбавленому пам'яті.

Для обчислення показника H відомий британський гідролог Х.Е. Херст запропонував метод нормованого розмаху (*R/S*-аналіз). Крім того, на основі *R/S*-аналізу обчислюється *R/S*-статистика для перевірки статистичної гіпотези про наявність довгострокової залежності. Розглянемо послідовність кроків обчислювального процесу за методом Херста [ ]:

1) Хай P_i (*i=1,2,…,N*) – ряд котирувань деякого цінного паперу (у загальному випадку – часовий ряд). Утворюємо з даного ряду послідовність $h_i = \ln\left(\dfrac{P_i}{P_{i-1}}\right)$ (*i=1,2,…,N-1*) - логарифм прибутковості.

2) Разоб'ем на l суміжних послідовностей $a_1, a_2, …, a_l$ однакової довжини n $l*n = N-1$. Позначимо h_{ik} елементи k-ой послідовності $k = 1,2,..l$.

3) Для кожної послідовності a_k обчислимо:

— середнє значення $\bar{h}k$ і стандартне відхилення S_k її елементів відповідно по формулах $\bar{h}k = \frac{1}{n}\sum_{i=1}^{n} h_{ik}$

$$S_k = \sqrt{\frac{1}{n}(h_{ik} - \bar{h})^2} ;$$

— накопичені суми $z_i = \sum_{r=1}^{i}\left(h_{rk} - \bar{h}_k\right)$ $i = 1,2,..,n$;

— розмах накопичених сум $R_k = \max_i z_i - \min_i z_i$, $i = 1,2,..,n$.

4) Для розбиття на підпослідовності довжини n обчислюємо нормований розмах накопичених сум $R/S = \frac{1}{l}\sum_{k=1}^{l}\frac{R_k}{S_k}$.

Далі проводять нове розбиття на підпослідовності більшої довжини і кроки 1-4 повторюють до $n = \left(M-1\right)/2$. Відзначимо, що невеликі значення n проводять нестабільні оцінки при невеликих об'ємах вибірки, тому рекомендується використовувати значення $n \ge 10$. Таким чином, буде набуто декілька значень *R/S* для різних *n*. Розмах *R/S* є відстанню, на яку переміщається система за час $T = n$. Херст виявив загальну форму рівняння

$$R/S = cn^H , \quad (1)$$

де c - константа. Значення *R/S* змінює масштаб у міру збільшення приросту часу *n* згідно значенню статечної функції, рівному n^H тобто діапазон збільшується згідно ступеню. Це називається масштабуванням із статечною залежністю, що є характерною межею фракталів. Логарифмуючи рівність (1), отримаємо рівняння прямої (2) в координатах ($\ln\left(R/S\right)\left[\ln n\right]$):

$$\ln\left(R/S\right) = \ln c + H \ln n \qquad (2)$$

R/S-аналіз є простим процесом, але він вимагає обробки великої кількості даних. Завдяки розвитку комп'ютерних технологій створені програми, за допомогою яких обчислюють коефіцієнт Херста. Одна з них Fractan 4.4, яка розповсюджується безкоштовно.

Аналіз і моделювання індексу ПФТС. Дослідження значень індексу ПФТС по роках з 2000 по 2010 (дані сайту www.pfts.ua) показали, що в цих рядах присутня стохастична складова. Про це свідчать оцінки спектральної щільності і кореляційної функції: періодограма і корелограма (Рис. 2 і 3). Аналогічний результат отриманий по даним за всі роки даного періоду. Таким чином, для моделювання і прогнозування потрібно використовувати дробноінтегровану модель.

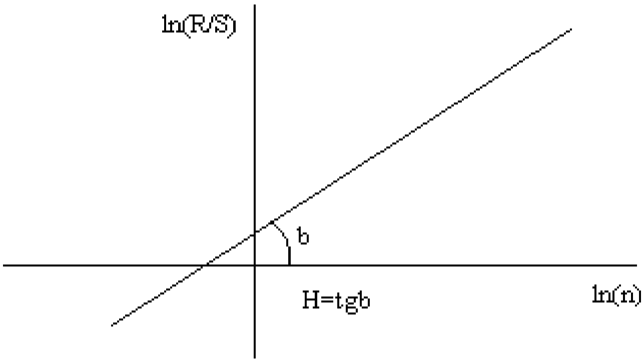

Рис. 1. Интепретація показателя Херста.

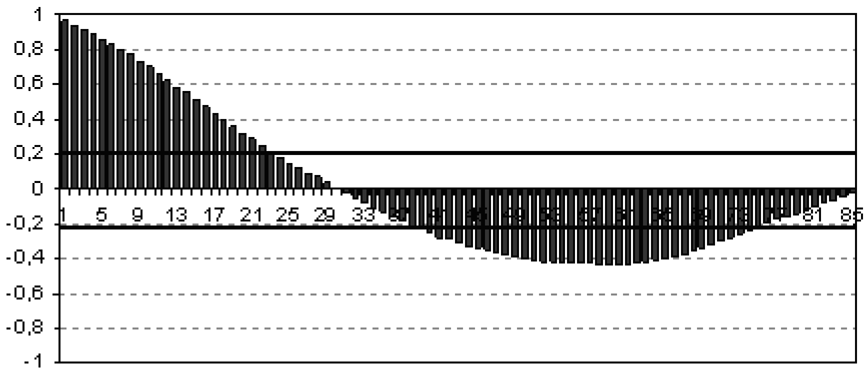

Рис. 2. Корелограма значень індексу ПФТС за 2010 р.

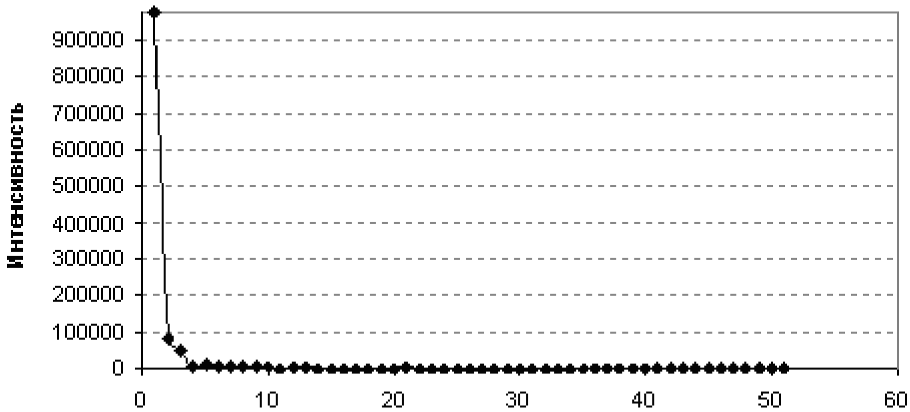

Рис. 3. Періодограмма значень індексу ПФТС за 2010р.

Моделювання в пакеті Fractan потрібне більше трьох тисяч емпіричних даних. В даний час, завдяки простоті практичного застосування, особливої популярності для аналізу і прогнозування тимчасових рядів набувають додатки – нейроімітатори, що використовують для аналізу і прогнозу нейронні мережі різної архітектури. Наприклад, вільно розповсюдженні NeuroPro, NeuroPro.25 і NeuroShell. Пакети містять програму, надбудову Excel, яка дає можливість застосовувати нейронні мережі з робочих листів Excel. Проте, застосування нейроімітаторів для вирішення практичних завдань є «скоріше мистецтвом, чим наукою» [8], оскільки вибір численних параметрів, які потрібні для моделювання, здійснюється тільки на основі особистого практичного досвіду розробника. Проте, на практиці для моделювання віддають перевагу використанню саме нейроімітаторам.

На рис. 3 показані результати моделювання і прогнозування на 20 днів за допомогою нейроімітатора NeuroPro 0.25. Чорним кольором позначені фактичні дані, сірим – прогнозні.

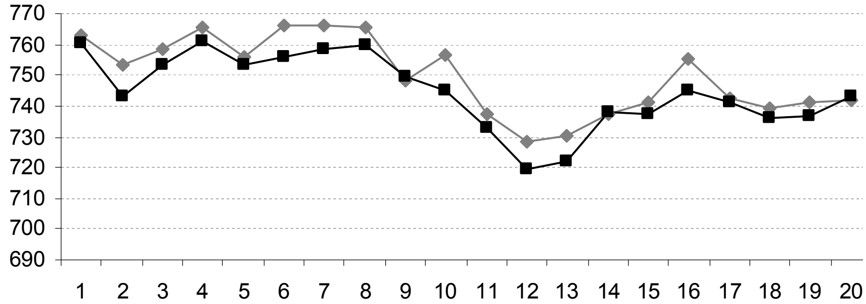

Рис. 3. Результат нейромоделювання

З малюнка видно, що модель повністю відображає тенденцію даних, при цьому середньоквадратична помилка $\sigma = 6.39$. Таким чином, можна вважати модель адекватною, а прогноз цілком задовільним.

Природно, встає проблема виявлення «довгої пам'яті», тобто для перевірки статистичної гіпотези про довгострокову (*long-range*) залежність. Для виявлення *long-range* залежності Б. Мандельброт запропонував використовувати *R/S*-статистику (*rescaled range*), розроблену Х. Херстом (1951). Це завдання вирішується наступнім шляхом:

Нехай гіпотеза H_0 - має місце короткострокова залежність, при альтернативній H_l - довгострокова залежність [7]. Хай є спостереження $X_1, X_2, \ldots, X_n$. Позначимо арифметичне середнє через $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Визначимо статистику

$$Q_n = \frac{1}{\sigma_n}\left[\max_{1\le k\le n}\sum_{j=1}^{k}(X_j - \overline{X}_n) - \min_{1\le k\le n}\sum_{j=1}^{k}(X_j - \overline{X}_n)\right]$$

де

$$\sigma_n^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2 + \frac{2}{n}\sum_{j=1}^{q}\varpi_j(q)\left(\sum_{i=j+1}^{n}(X_i - \overline{X}_n)(X_{i-j} - \overline{X}_n)\right),$$

$$\varpi_j(q) = 1 - \frac{j}{q+1}, q < n$$

По суті статистика Q_n є нормований розмах *R/S* накопичених сум, тому часто цю статистику називають *R/S* тестом або *R/S* статистикою . Слід зазначити, що у відмінності від *R/S*, оцінка дисперсії σ_n^2 відображає не тільки суму квадратів відхилень, але і зважені автоковаріації до деякого лага *q*. Ваги $\varpi_j(q)$ забезпечують додатність σ_n^2.

Проте дуже великі (в порівнянні з *n*) значення *q* приводять до того, що властивості оцінок в скінченних вибірках істотно відрізнятимуться від їх асимптотичної поведінки. Але брати дуже маленькі значення *q* теж не можна, оскільки автоковаріації за лагом *q* можуть виявитись значними і їх слід було б включити в зважену суму. Д. Ендрюс запропонував для вибору *q* правило, засноване на властивостях початкових даних [9]. В якості *q* беруть цілу частину числа

$$k_n = \left(\frac{3n}{2}\right)^{1/3}\left(\frac{2\rho}{1-\rho^2}\right)^{2/3},$$

де ρ є оцінкою коефіцієнта автокореляції першого порядку X_t.

Асимптотичний розподіл статистики $V_n(q) = \frac{1}{\sqrt{n}}Q_n$ збігається до розподілу випадкової величини - розмаху броунівського моста на одиничному інтервалі [7]. Нижче наведена таблиця критичних значень $V_n(q)$.

Таблиця 2.7 Критичні значення статистики $V_n(q)$

0,5%	(0,721; 2,098)
2,5%	(0,089; 1,862)
5%	(0,861; 1,747)
10%	(0,927;1,620)

Таблиця 2.7 містить інтервали, при попаданні в які статистики V_n нульова гіпотеза про коротко-строкову залежність приймається.

Для кількості спостережень $n = 252$ індексу ПФТС отримані наступні значення: $\rho = 0{,}97$, $k_n = 72.51$, q=72, $\sigma = 2.58$, $Q_n = 49.6992$, $V_n = 3.1307$. Значення статистики V_n не потрапляє ні в один критичний інтервал. Отже, гіпотеза відхиляється, тобто немає підстав вважати, що досліджуваний ряд володіє короткостроковою залежністю. Таким чином, индекс ПФТС має довгострокову пам'ять. Коефіцієнт Херста $H = 0.7694$, що вказує на персістентні властивості індексу.

Наявність стохастичного тренда і персисиентної властивості свідчить про інерційність фондового ринку, що дає можливість під іншим кутом поглянути на роботу фінансових ринків і методологію оцінки ризиків.

Список літератури:

1. Wold H. A Study in the Analysis of Stationary Time Series. / H. Wold // Сб.научн.трудов. Stockholm: Almqvist and Wiksel, 1938.
2. Канторович Г.Г. Анализ временных рядов / Г.Г. Канторович // Экономический журнал ВШЭ. – 2002. – №2. – С. 252-273.
3. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс. – М.:Мир, 1974. – 197 с.
4. Granger C.W.J. Some Properties of Time Series Data and Their Use in Econometric Model Specification / C.W.J. Granger // Journal of Econometrics. – 1981. – Vol.16, №1. – Р3. 121-130.
5. Леоненко М.М. Теоретико-ймовірнісні та статистичні методи в економетриці та фінансовій математиці / М.М. Леоненко, Ю.С. Мішура, В.М. Пархоменко, М.Й. Ядренко - К.: Інформтехніка, 1995. – 380 с.
6. Петерс Э. Фрактальный анализ финансовых рисков / Э. Петерс. – М.:2004. – 304 с.
7. Lo A.W. Long Term Memory in Stock Market Prices / A.W. Lo // Econometrica. – 1991. – №59. – PP.1279-1313.
8. Хайкин С. Нейронные сети: полный курс, 2-е изд., испр./ Хайкин С. Пер. с англ. – М.: ООО "И.Д. Вильямс", 2006. - 1104 с./ Под ред. д.т.н. Н.Н. Куссуль.
9. Andrews D. Non-Strong Mixing Autoregressive Process//Journal of Probability.-1984.-№21.-P.-930-934.

Надано до редакції 20.11.2011

Андрієнко Валентина Михайлівна / Valentina M. Andrienko
andrienko.v@gmail.com

Співаков Олег Геннадійович / Oleg G. Spivakov
o.spivakov@gmail.com