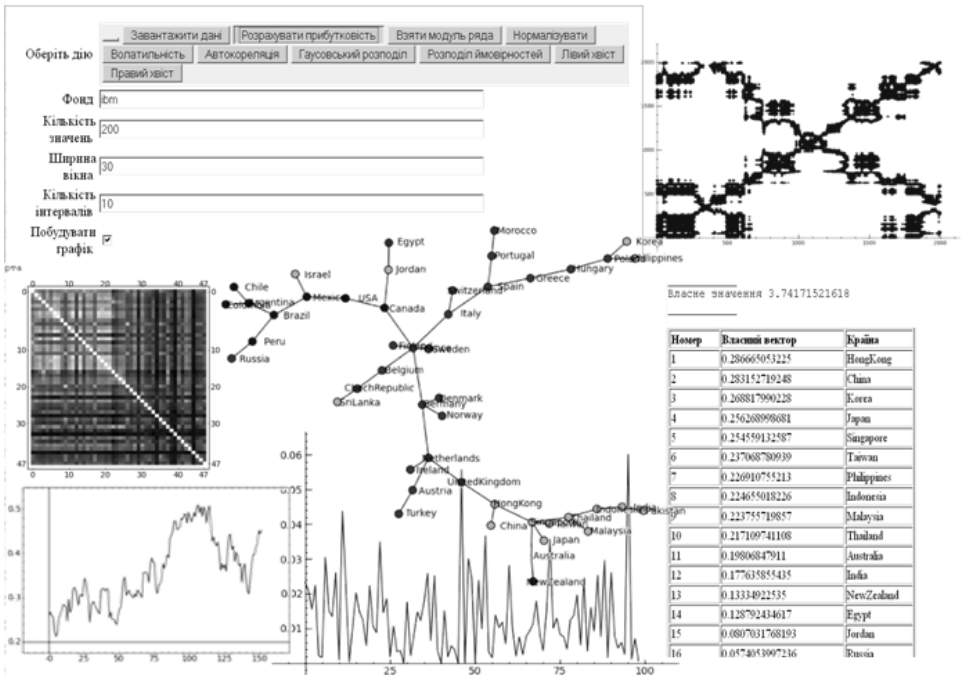


# Моделювання складних економічних систем

*навчальний посібник*



*В. М. Соловйов*  
*В. В. Соловйова*  
*Н. А. Хараджян*

# Моделювання складних економічних систем

*навчальний посібник*

Кривий Ріг  
Видавничий відділ НМетАУ  
2010



УДК 378.147+330.46:004.9432

Соловійов В. М. Моделювання складних економічних систем : навчальний посібник / В. М. Соловійов, В. В. Соловійова, Н. А. Хараджян. – Кривий Ріг : Видавничий відділ НМетАУ, 2010. – 119 с.

Посібник призначений для вивчення спецкурсу «Моделювання складних економічних систем» у процесі підготовки фахівців з економічної кібернетики засобами комп'ютерного моделювання.

В якості середовища моделювання обрано Web-СКМ Sage. Посібник містить необхідний теоретичний матеріал, що супроводжується лабораторними роботами та індивідуальними завданнями.

Для студентів спеціальності «Економічна кібернетика», аспірантів та викладачів.

Рецензенти:

**І. І. Кукурудза**, доктор економічних наук, професор, завідувач кафедри економічної теорії та міжнародної економіки Черкаського національного університету імені Богдана Хмельницького;

**Ю. В. Триус**, доктор педагогічних наук, професор, професор кафедри комп'ютерних технологій Черкаського державного технологічного університету

Друкується згідно з рішенням вченої ради Черкаського національного університету імені Богдана Хмельницького (протокол № 2 від 16 листопада 2010 р.)

© В. М. Соловійов, В. В. Соловійова, Н. А. Хараджян, 2010

## ЗМІСТ

<b>ВСТУП.....</b>	<b>5</b>
<b>1. ПОЧАТОК РОБОТИ У SAGE.....</b>	<b>8</b>
<b>2. АНАЛІЗ ФЛУКТУАЦІЙ ЧАСОВОГО РЯДУ.....</b>	<b>15</b>
2.1. Аналіз динаміки прибутків, модулів прибутків та волатильностей 15	
2.1.1. Визначення волатильності.....	16
2.1.2. Визначення кореляцій.....	17
2.2. Хід виконання роботи.....	18
2.3. Завдання для самостійної роботи.....	23
2.4. Контрольні питання.....	24
<b>3. МУЛЬТИФРАКТАЛЬНИЙ АНАЛІЗ ФЛУКТУАЦІЙ ЧАСОВОГО РЯДУ.....</b>	<b>25</b>
3.1. R/S-аналіз.....	25
3.2. Стандартний аналіз флуктуацій.....	26
3.3. Аналіз детрендованих флуктуацій.....	28
3.3.1. Локальний АДФ.....	30
3.3.2. Мультифрактальний АДФ.....	30
3.4. Хід виконання роботи.....	33
3.5. Завдання для самостійної роботи.....	42
3.6. Контрольні питання.....	42
<b>4. ДОСЛІДЖЕННЯ ПРОЦЕСІВ САМООРГАНІЗАЦІЇ В ЕКОНОМІЧНИХ СИСТЕМАХ.....</b>	<b>43</b>
4.1. Моделювання колективних ефектів складних систем засобами теорії випадкових матриць.....	43
4.1.1. Знаходження коефіцієнтів матриці крос-кореляцій.....	44
4.1.2. Розподіл власних значень.....	45
4.1.3. Обернене відношення участі.....	47
4.2. Інтерпретація власних значень та власних векторів, що відхиля- ються від ТВМ.....	47
4.3. Хід виконання роботи.....	48
4.4. Завдання для самостійної роботи.....	55
4.5. Контрольні питання.....	55
<b>5. КЛАСТЕРНИЙ АНАЛІЗ.....</b>	<b>56</b>
5.1. Кластерний аналіз на основі часових рядів.....	56
5.1.1. Гіпотеза ефективного ринку.....	56
5.1.2. Отримання даних з часової послідовності.....	56
5.2. Побудова матриці відстаней на основі матриці крос-кореляцій.....	57
5.3. Побудова мінімального остівного та ієрархічного дерев.....	59
5.4. Хід виконання роботи.....	60

5.5. Завдання для самостійної роботи.....	62
5.6. Контрольні питання.....	62
<b>6. ДОСЛІДЖЕННЯ ДИНАМІКИ СКЛАДНИХ СИСТЕМ ЗА ХАРАКТЕРИСТИКАМИ КЛАСТЕРНОГО ДЕРЕВА.....</b>	<b>64</b>
6.1. Коефіцієнт кластеризації графа.....	64
6.2. Середня довжина шляху на графі.....	65
6.3. Діаметр графа.....	66
6.4. Максимальна степінь вершин графа.....	66
6.5. Хід виконання роботи.....	67
6.6. Завдання для самостійної роботи.....	70
6.7. Контрольні питання.....	70
<b>7. ВИКОРИСТАННЯ РЕКУРЕНТНОГО АНАЛІЗУ ДЛЯ МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ НЕЛІНІЙНИХ ДИНАМІЧНИХ ВЛАСТИВОСТЕЙ СКЛАДНИХ СИСТЕМ.....</b>	<b>71</b>
7.1. Застосування рекурентного аналізу та рекурентних діаграм до дослідження динаміки та топології складних систем.....	71
7.2. Рекурентний аналіз.....	73
7.3. Аналіз діаграм.....	79
7.4. Хід виконання роботи.....	81
7.5. Завдання для самостійної роботи.....	85
7.6. Контрольні питання.....	85
<b>8. КІЛЬКІСНИЙ АНАЛІЗ РЕКУРЕНТНИХ ДІАГРАМ.....</b>	<b>86</b>
8.1. Структури в рекурентних графіках.....	86
8.2. Хід виконання роботи.....	88
8.3. Завдання для самостійної роботи.....	91
8.4. Контрольні питання.....	92
<b>ДОДАТКИ.....</b>	<b>93</b>
А. Модуль finance.....	93
Б. Модуль Graph.....	103
В. Модуль dynamics.....	109
<b>ЛІТЕРАТУРА.....</b>	<b>116</b>

## Вступ

Останні роки стало зрозуміло, що соціально-економічні системи відносяться до класу так званих складних систем, ефективне дослідження яких можливе засобами виникаючої у нас на очах міждисциплінарної науки – теорії складних систем [8.4]. Оскільки подібні системи практично виключають можливості їх аналітичного дослідження, найбільш дієвим інструментарієм стають методи комп'ютерного моделювання [8.4; 8.4]. Фізиками накопичено потужний арсенал різноманітних моделей, адекватне перенесення яких на об'єкти нефізичної природи дозволяє краще зрозуміти структурні та динамічні аспекти функціонування досліджуваних складних систем. Тому не дивно, що з кінця минулого століття фізики все частіше стали приділяти увагу фінансовим та економічним системам, використовуючи інструменти та методологію, специфічну для фізичних наук. Такий вибір був мотивований кількома причинами. Перш за все, деякі фінансові та економічні системи є «гарними» екземплярами складних систем. По-друге, починаючи із 80-х років минулого століття велика кількість фінансових та економічних даних була записана у комп'ютерній формі і є легкодоступною для аналізу та тестування. Такі дані відображають детальну інформацію про процеси, що розглядаються, і у випадку деяких бірж, містять інформацію про елементарні взаємодії, що відбуваються на біржі (існують бази даних торгів та квот, де кожна транзакція, що виконується на біржі, документується). По-третє, аналіз та моделювання економічних і фінансових систем має базові теоретичні та прикладні аспекти. Базові аспекти стосуються моделювання систем, скомпонованих із кількох складових (підсистем), взаємодія яких зазвичай описується нелінійними моделями. Практичні аспекти стосуються жорсткого визначення та кращого виявлення притаманного системі економічного ризику. Нарешті, фінансово-економічні системи мають помітне прагматичне значення. Це яскраво продемонструвала глобальна фінансова криза 2007-2009 рр., яка торкнулася майже кожного. Стало зрозуміло, що моніторинг, прогнозування та попередження подібних кризових явищ стає надзвичайно актуальною задачею.

На поточний момент опубліковано досить велику кількість робіт, що стосуються розгляданого матеріалу, у фізичних, математичних та економічних журналах. Найбільш часто у статтях розглядається кілька питань:

- вивчення статистичних властивостей зміни ціни акцій;
- новітні підходи до проблеми ціни опціонів;
- дослідження моделей фінансових ринків;
- порівняння між динамікою цін на фінансових ринках та швидкість зміни динаміки.

Приклади, наведені вище, не вичерпують весь спектр дослідницьких проблем, і є інші питання, які вимагають дослідження.

Традиційні аналітичні методи дослідження економічних, фінансових, соціальних систем все частіше наштовхуються на проблеми, що не мають ефективного розв'язання в рамках класичних парадигм. Класичні підходи були розроблені для опису сталого світу, який поволі еволюціонує. За самою своєю суттю ці методи і підходи не були призначені для опису та моделювання швидких змін, непередбачуваних стрибків і складних взаємодій окремих складових сучасного світового ринку [8.4].

У процесі розробки нових парадигм досліджень складних економічних систем виникали численні прийоми та підходи, що поступово нагромаджувались, удосконалювались, узагальнювались, закладаючи основи технології та методології подолання зазначених утруднень, сприяючи розвитку таких міждисциплінарних наукових напрямів, як теорія систем, системний аналіз і кібернетика.

Економічна кібернетика тісно пов'язана, з одного боку, з теорією управління, сучасними інформаційними системами та технологіями, а з другого – з багатьма конкретними економічними дисциплінами (економічною теорією, макро- та мікроекономікою, менеджментом та ін.). Спираючись на поєднання цих наук, економічна кібернетика формує цілісне уявлення про економіку як складну динамічну систему, вивчає взаємодію виробничо-технічної, соціально-економічної та організаційно-господарської структур економіки у процесі управління нею, а також у процесах функціонування й розвитку цієї системи як єдиного цілого. При цьому фундаментальною основою економічної кібернетики виступає теорія систем, а провідним методом дослідження – метод моделювання.

Протягом останніх років відбулися суттєві зміни в розумінні фундаментальних закономірностей економічних систем. Було виявлено, що такі системи мають універсальні емерджентні властивості, які не знаходять адекватного розуміння в рамках традиційних парадигм. Тому для дослідження економічних систем все активніше використовуються методи та моделі фундаментальних наук у поєднанні з

сучасними інформаційними технологіями. Так, застосування методології фізики до вивчення економічних систем привело до появи нового розділу математичної економіки – еконофізики [8.4].

Поява нових методологій досліджень підвищує вимоги до професійної компетентності та фахової мобільності фахівців, що в свою призводить до появи нових курсів та спецкурсів, зокрема спецкурсу «Моделювання складних економічних систем», здатного готувати сучасних фахівців, які вміють аналізувати мінливі соціально-економічні тенденції, приймати і реалізовувати нестандартні рішення в ситуації ринкової конкуренції.

Спецкурс передбачає оволодіння теоретичними та практичними основами моделювання складних економічних систем; різними характеристиками економіки (емерджентність, динамічність, невизначеність), проблеми методології макроекономічного аналізу. Значна увага приділяється моделюванню критичних, кризових явищ та універсальності складних систем, розглядається дослідження процесів самоорганізації в економічних системах. Передбачається ознайомлення з теорією випадкових матриць, що була розвинена для пояснення статистики рівнів енергії складних квантових систем. Наводяться результати моделювання колективних ефектів складних систем за допомогою методології теорії випадкових матриць. Значна увага приділяється розгляду методів дослідження динаміки і топології фінансово-економічних систем за допомогою фрактального аналізу, пошуку та конструювання індикаторів передкризових станів за допомогою аналізу фрактальних та рекурентних властивостей складних динамічних систем. Завершує спецкурс ознайомленням із використанням рекурентного аналізу для моделювання і прогнозування нелінійних динамічних властивостей складних систем та з інструментарієм нелінійної динаміки, що відноситься до рекурентних властивостей нестационарних динамічних рядів.

Сьогодні здійснення розрахунків, проведення дослідження та моделювання явищ різної природи у різних предметних галузях засобами систем комп'ютерної математики (СКМ) є вже традиційними і не потребує обґрунтування. Але при моделюванні складних економічних систем бажано використовувати таке середовище, що є здатним інтегрувати в собі інші СКМ, будучи вільнопоширюваним та оснащеним Web-інтерфейсом. Порівняльний аналіз існуючих Web-СКМ показують доцільність вибору в якості середовища моделювання Web-СКМ Sage.

## 1. Початок роботи у SAGE

Sage (<http://sagemath.org>) являється прикладом безкоштовного вільнопоширюваного середовища математичних обчислень для виконання символьних, алгебраїчних та чисельних розрахунків, інтерфейс та ядро якого написані потужною і досить популярною мовою програмування Python. Sage об'єднує можливості популярних математичних програм та бібліотек, таких як PARI, GAP, GSL, Singular, MWRANK, NetworkX, Maxima, Sympy, GMP, Numpy, matplotlib та багатьох інших засобів: Python, Lisp, Fortran 95, C/C++ тощо.

Sage є серверним програмним засобом, що базується на відомому об'єктно-орієнтованому сервері додатків Python-CMS Zope. Інтерактивність Web-клієнта забезпечується широким застосуванням технології AJAX, що є основою більшості продуктів Web 2.0, а відповідність відображення математичної інформації – браузерними математичними шрифтами (jsMath).

Sage має власне символьне ядро, проте виступає переважно як інтегратор різних систем, надаючи їм єдиний Web-інтерфейс. Можливість виконання на Web-сторінках, генерованих Sage, програм мовами Fortran, Python, Lisp, Java та ін., надає їм надвисокого рівня інтерактивності, порівняного з комерційними СКМ (Mathematica, Matlab, MathCAD, Maple), без суттєвих вимог до апаратних ресурсів комп'ютера користувача (необхідні лише браузер та мережне з'єднання).

Основними складовими Sage є:

– інтерфейси до СКМ Magma, Maple, Mathematica, Matlab, MuPAD та ін.;

– якісні пакети для алгебри та обчислень (Maxima), швидких високоточних обчислень (GMP), алгебраїчної геометрії (Singular), лінійної алгебри (Linbox), графіки (Gnuplot), теорії груп (GAP), теорії чисел (PARI), оптимізації (GSL) та ін.

– мови програмування (Python, Lisp, Fortran 95, C/C++ та ін.).

Sage має Web-інтерфейс (рис. 1.1).

Web-інтерфейс Sage отримав назву блокнотного (*notebook*), тому що він являє собою комп'ютерну модель записника, який математики традиційно використовують для виконання математичних розрахунків. Подібний інтерфейс є традиційним для розвинених СКМ.

Наявність Web-інтерфейсу, безкоштовність та відкритість

середовища математичних обчислень Sage – це основні, але не єдині переваги програми у порівнянні з популярними системами комп’ютерної математики. Слід додати такі характеристики Sage:

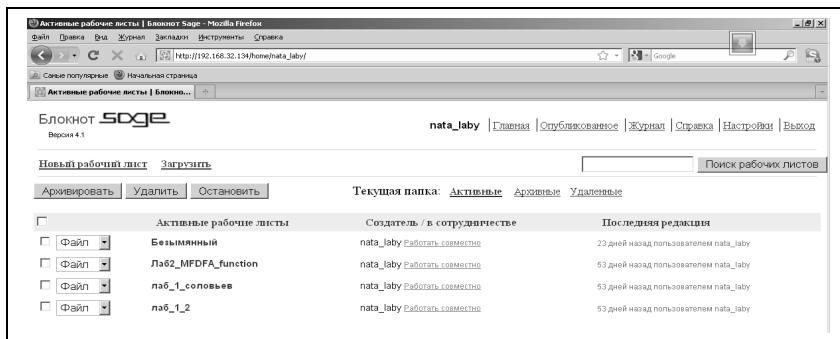


Рис. 1.1 Web-інтерфейс Sage

– невимогливість до апаратної складової обчислювальної системи;

– індиферентність до використовуваного браузера;

– для організації роботи у мережі достатньо встановити Sage на сервері;

– підтримка інтерфейсів до комерційних систем комп’ютерної математики, таких як Maple, Magma, Mathematica і Matlab;

– для подання математичних виразів у звичній математичній нотації немає потреби встановлювати спеціальне програмне забезпечення – достатньо дозавантажити математичні шрифти;

– наявність потужного інструментарію для побудови статичних та динамічних графічних зображень (на площині та у просторі);

– максимальне спрощення процедури публікації робочих аркушів записника Sage у мережі Інтернет;

– підтримка засобів спільної роботи.

Для організації роботи у локальній мережі достатньо встановити Sage на будь-якому комп’ютері.

З метою формування первинних навичок проведення моделювання у новому програмному середовищі до запропонованих тем з моделювання складних економічних систем, було включено тему «Початок роботи у Sage», що входить до посібника [8.4]. Даний посібник можна використовувати як для самостійного опанування Sage, так і в якості довідника. У додатках до посібника наведені розвинені



прикладі застосування Sage для побудови інтерактивних моделей та розв’язання задач теорії кодування. В останньому додатку показано, як можна розширити можливості Sage та інтегрувати у нього нове програмне забезпечення.

Після виконання всіх операцій з підготовки Sage до роботи та його запуску за виділеною IP-адресою у вікні Web-браузера буде відкрита домашня сторінка користувача Sage (рис. 1.2).

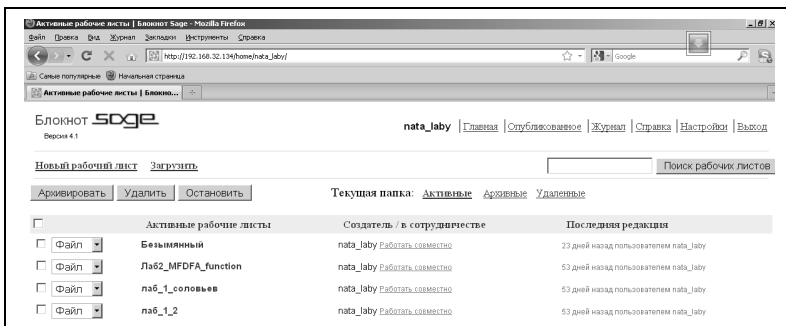


Рис. 1.2. Домашня сторінка користувача Sage (Worksheet List)

На домашній сторінці представлений персональний список робочих листів. Робочі листи (*Worksheets*) – це основні елементи Sage, оскільки математичні розрахунки виконуються саме на робочих листах.

Всі робочі листи блокноту розподілені по папках за категоріями:

- активні листи (листи, необхідні користувачу для роботи найближчим часом);
- архівні листи;
- листи, підготовлені до вилучення.

На початку сеансу роботи з блокнотом Sage поточною папкою (*Текущая папка*) є папка з активними листами – *Активные*. Додати до списку перегляду архівні листи можна, обравши посилання *Архивные*, а перейти до режиму перегляду листів, підготовлених до вилучення – посилання *Удаленные*.

Активні листи можна перемістити до папки-архіву (кнопка *Архивировать*) або до папки-кошика (кнопка *Удалить*). Листи з архіву можуть бути перенесені до папки *Активные* (кнопка *Архивировать*) або вилучені. У кошику Sage з робочими листами можуть бути виконані дії, аналогічні до тих, які передбачені у Кошику операційної системи Windows – остаточного знищення або відновлення вилуче-

них листів (*Восстановить*). Перш ніж дати команду перенесення листа (чи листів) між папками, їх треба відмітити пташкою.

Над робочими листами, як над файлами, можна виконати операції перейменування (*Переименовать*), редагування (*Редактировать*), копіювання (*Копировать рабочий лист*), публікації в мережі Інтернет (*Опубликовать*), перегляду ходу роботи з листом (*Изменения*). Окрім того, можна визначити перелік користувачів, які можуть спільно працювати з певним робочим листом (*В сотрудничестве*). Список зазначених дій представлений на домашній сторінці користувача у вигляді випадаючого списку *Файл* ліворуч від заголовку кожного робочого листа.

Вибрати робочий лист для подальшої роботи користувач може одним із зазначених способів:

1) створити новий робочий лист, вибравши посилання *Новый рабочий лист*;

2) завантажити робочий лист (файл з розширенням *sws*) з довільного носія або з Інтернету за вказаною URL-адресою, вибравши посилання *Загрузить*;

3) відкрити існуючий робочий лист, вибравши його зі списку. У разі завеликої кількості листів передбачена можливість здійснення автоматичного пошуку необхідного робочого листа (*Поиск Рабочего листа*), аналогічно до пошуку звичайної Web-сторінки в межах сайту;

4) завантажити всі активні листи можна за допомогою *Загрузить все активные*.

Web-сторінка з робочим листом та його основними елементами представлені на рис. 1.3.

З даними робочого листа можна працювати у режимах *Робочий лист*, *Редактировать* та *Текст*.

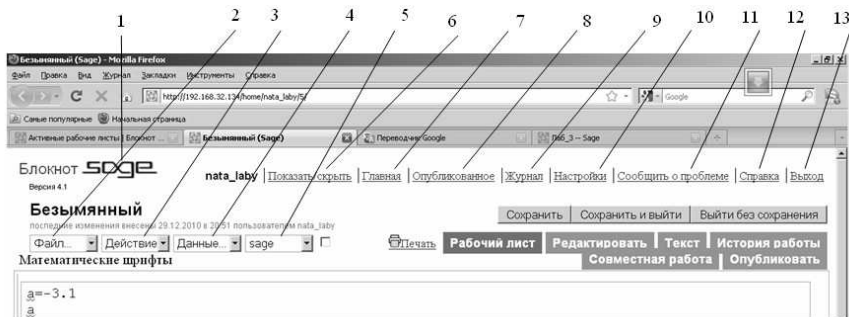
Режим *Робочий лист* – основний режим роботи – режим перегляду та виконання математичних розрахунків.

Режим *Редактировать* – режим редагування. У даному режимі користувач має можливість додати до змісту робочого листа тексти завдань та виконати різні маніпуляції з блоками математичних розрахунків (вилучати, переміщувати та ін.).

*Зауваження:* завершуючи роботу в цьому режимі, обов'язково збережіть зміни (кнопка *Сохранить изменения*), інакше вони будуть відкинуті без попередження.

Режим *Текст* служить для відображення даних робочого листа

(завдань, команд та результатів) у текстовому (нерозміченому) форматі.



- 1 – банер-посилання на сайт `www.sagemath.org`;
- 2 – список *Файл* допустимих операцій над листами блокноту;
- 3 – список *Действие* допустимих дій над даними листа;
- 4 – список *Данные* для роботи з файлами даних;
- 5 – список пакетів, інтерфейси яких підтримує Sage;
- 6 – *Показать скрытые* – приховати/показати панель управління;
- 7 – *Главная* – перейти до персонального списку листів;
- 8 – *Опубликованные* – показати список опублікованих листів;
- 9 – *Журнал* – показати історію останніх дій;
- 10 – *Настройки* – параметри облікового запису;
- 11 – *Сообщить о проблеме* – повідомити про проблему або описати помилку для покращення роботи Sage;
- 12 – *Справка* – викликати довідкову систему;
- 11 – *Выход* – завершити сеанс роботи з блокнотом.

Рис. 1.3. Вікно робочого листа

Вибір *История работы* дозволяє переглянути історію роботи над листом.

За допомогою *Совместная работа* визначається перелік користувачів, які можуть спільно працювати з певним робочим листом.

*Опубликовать* відкріє вікно для підготовки робочого листа до публікації у мережі.

В основному режимі роботи користувач має можливість ввести у прямокутну комірку команду чи програмний код для виконання математичних розрахунків та переглянути результат обчислень.

Автодоповнення вводу в Sage суттєво спрощує процес введення команд: увівши перші літери команди і натиснувши клавішу табуля-

ції, користувач отримує список команд, що починаються з введеної послідовності символів.

Дати команду «розпочати обчислення» за введеною командою чи програмним кодом у комірці можна, натиснувши комбінацію клавіш `<Shift> + <Enter>` або обравши посилання `evaluate` під командним рядком. Під коміркою з'явиться зелена вертикальна риска, яка вказує на те, що на сервері Sage виконуються обчислення. По завершенню обчислень риска зникає і під коміркою з'являється результат обчислення або повідомлення про допущену помилку.

Результати обчислень можуть бути подані у текстовому форматі (за замовчуванням) або у природній математичній формі. Для відображення математичних виразів у природній формі слід встановити математичні шрифти викликом *jsMath Control Panel* у відповідному вікні повідомлення (рис. 1.4).

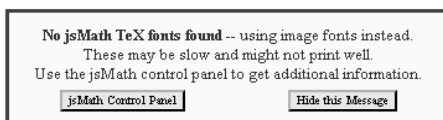


Рис. 1.4. Вікно попередження jsMath

Найчастіше помилки виникають через неправильний синтаксис команди. Для виправлення таких помилок доречним є звернення до довідкової системи Sage. Контекстну довідку можна отримати, ввівши символ `?` після імені команди та `<Shift> + <Enter>`.

Введення символу `??` після назви функції дозволяє переглянути її програмний код.

Після виведеного результату на робочому листі автоматично з'являється нова комірка для введення команди. Додати нову комірку у довільному місці робочого листа можна вибором горизонтальної синьої лінії, яка з'являється у момент нависання курсору миші над будь-якою коміркою. Вилучення зайвої порожньої комірки виконується клавішею `<Backspace>`.

За замовчуванням робочий лист налаштований на роботу з командами Sage. Вибрати іншу СКМ можна у списку (5) або командою виду `%назва_СКМ`. Наприклад: `%maxima`, `%maple` і т.д.

Основні дії над робочим листом представлені у списку *Действие...* (3):

*Прервать* – перервати процес обчислень.

*Перезапустити робочий лист* – перезапустити сеанс роботи з робочим листом.

*Сохранить робочий лист и вийти* – завершити сеанс роботи з робочим листом.

*Вычислить все* – виконати програмні коди у всіх комірках.

*Скрыть весь вывод* – приховати всі результати виконання команд.

*Показать весь вывод* – показати всі результати виконання команд.

*Удалить весь вывод* – видалити результати виконання всіх комірок.

*Одноячеечный режим* – залишити доступною для роботи лише одну першу комірку листа.

*Многаячеечный режим* – зробити доступними для роботи всі комірки листа.

Для завантаження чи створення файлу даних передбачена відповідна команда *Загрузить или создать файл* списку *Данные...* (4).

Завершити сеанс роботи з листом можна одним із передбачених способів, а саме:

– зберегти зміни – кнопка *Сохранить*;

– зберегти зміни та закрити вікно з робочим листом – кнопка *Сохранить и вийти*;

– проігнорувати зміни та закрити вікно з робочим листом – кнопка *Вийти без сохранения*.

При виконанні лабораторних робіт необхідно завантажити підготовлені моделі із папки, яку вкаже викладач. Після завантаження файлу з моделлю необхідно виконати всі комірки, обравши пункт *Выполнить все*. Після виконання всіх комірок з'явиться вікно моделі де виконуємо необхідні дії, що описані в кожній лабораторній роботі.

## 2. АНАЛІЗ ФЛУКТУАЦІЙ ЧАСОВОГО РЯДУ

### 2.1. Аналіз динаміки прибутків, модулів прибутків та волатильностей

Останнім часом вчені все більше цікавляться економічними часовими рядами, і відбувається це за кількох причин, зокрема:

1) економічні часові ряди, такі як індекси акцій, курсів валют, залежать від розвитку великої кількості взаємодіючих чинників, і є прикладами складних систем, що широко вивчаються у науці;

2) з'явилась велика кількість доступних баз з даними про економічні системи, що містять інформацію з різними часовими шкалами (починаючи з 1 хвилини і закінчуючи 1 роком).

Внаслідок цього вже на даний час існує велика кількість розроблених методів (зокрема, у статистичній фізиці), спрямованих на отримання характеристик цін акцій чи курсів валют, що еволюціонують у часі.

Дослідження, проведені над часовими рядами, показують, що стохастичний процес, який лежить у основі зміни ціни, характеризується кількома ознаками. Розподіл зміни ціни має виділений хвіст порівняно із нормальним розподілом. Функція автокореляції зміни ціни спадає експоненційно з певним характерним часом. Однак, виявляється, що амплітуда зміни ціни, виміряна за абсолютними значеннями чи квадратами цін, показує степеневі кореляції з довгочасовою персистентністю аж до кількох місяців, або навіть років. Такі довго часові залежності краще моделюються з використанням «додаткового процесу», що в економічній літературі часто називається волатильністю. Волатильність змін ціни акції є мірою того, як сильно ринок схильний до флуктуацій, тобто відхилень ціни від попередніх значень.

Першим кроком при проведенні аналізу є побудова оцінювача волатильності. Ми будемо отримувати волатильність як локальне середнє модуля зміни ціни.

Розуміння статистичних властивостей волатильності має також важливе практичне застосування. Волатильність є інтересом торговців, оскільки визначає ризик і є ключовим входом практично до всіх моделей цін опціонів (вторинного цінного паперу), включаючи і класичну модель Блека-Шоулза. Без задовільних методів оцінювання волатильності трейдерам було б надзвичайно важко визначати

ситуації, в яких опціони попадають в недооцінку чи переоцінку.

### 2.1.1. Визначення волатильності

Термін волатильність представляє узагальнену міру величини ринкових флуктуацій (відхилень). В літературі існує досить багато визначень волатильності, проте ми будемо використовувати наступне: *волатильність є локальним середнім модуля зміни ціни на відповідному часовому інтервалі  $T$ , що є рухомих параметром нашої оцінки.*

Для індексу  $X(t)$  визначимо зміну ціни  $G(t)$  як зміну логарифмів індексів,

$$G(t) = \ln X(t + \Delta t) - \ln X(t) \cong \frac{X(t + \Delta t) - X(t)}{X(t)}, \quad (2.1)$$

де  $\Delta t$  є часовим інтервалом затримки. Величину (2.1) називають прибутковістю («*return*»). Якщо використовувати границі, то малі зміни  $X(t)$  приблизно відповідають змінам, визначеним другою рівністю. Ми лише підраховуємо час роботи ринку, викидаємо ночі, вихідні та свята із набору даних, тобто, вважається, що ринок працює без перерв.

Модуль  $G(t)$  описує амплітуду флуктуацій. У порівнянні із значеннями  $G(t)$  їх модуль не показує глобальних трендів, але великі значення  $|G(t)|$  відповідають крахам та великим миттєвим змінам на ринках [8.4].

Визначимо волатильність як середнє від  $|G(t)|$  для часових вікон  $T = n \cdot \Delta t$ , тобто:

$$V_T(t) = \frac{1}{n} \sum_{t'=t}^{t+n-1} |G(t')|, \quad (2.2)$$

де  $n$  є цілим числом. Таке визначення може бути ще узагальнене заміною  $|G(t)|$  на  $|G(t)|^\gamma$ , де  $\gamma > 1$  дає більш виражені великі значення  $|G(t)|$ , в той час як  $0 < \gamma < 1$  виділяє малі значення  $|G(t)|$ .

У цьому визначенні волатильності використовується два параметри,  $\Delta t$  та  $n$ . Параметр  $\Delta t$  є шаблонним (чи модельним) часовим інтервалом для даних, а параметр  $n$  є кроком переміщення часового вікна. Зауважимо, що вказане визначення волатильності має внутрішню помилку, а саме: вибір більшого часового інтервалу  $T$  веде до збільшення точності визначення волатильності. Однак, вели-

ке значення  $T$  також включає таке розбиття часу на інтервали, що веде до врахування не всієї прихованої у ряді інформації.

При порівнянні між собою кількох акцій часто використовують нормалізовану волатильність, що визначається для кожної акції наступним чином:

$$v_T^i = v_T^i(t) = \frac{V_T^i}{\sqrt{\langle (V_T^i)^2 \rangle - \langle V_T^i \rangle^2}}, \quad (2.3)$$

де  $\langle \dots \rangle$  означає середній час, отриманий для вікон, що не перекриваються, для різних часових масштабів  $T$ .

### 2.1.2. Визначення кореляцій

Для визначення кореляцій часового ряду використовується функція автокореляції [8.4]. Саме поняття *автокореляції* означає кореляцію часового ряду самого з собою (між попередніми та наступними значеннями). Автокореляцію іноді називають *послідовною кореляцією*, що означає кореляцію між членами ряду чисел, розташованих у певному порядку. Також синонімами цього терміну є *лагова кореляція* та *персистентність*. Наприклад, часто зустрічається автокореляція геофізичних процесів, що означає перенесення залишкового процесу на наступні часові проміжки.

Позитивно автокорельований часовий ряд часто називають персистентним, що значить існування тенденції слідування великих значень за великими та малих за малими, інакше позитивно корельований часовий ряд можна назвати інертним.

Візьмемо  $N$  пар спостережень двох змінних  $x$  та  $y$ . Кореляційний коефіцієнт між  $x$  та  $y$  визначається як:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}, \quad (2.4)$$

де сума знаходиться по всім  $N$  спостереженням.

Таким же чином можна визначити й автокореляцію, або ж кореляцію всередині одного часового ряду. Для автокореляції першого порядку береться лаг (часова затримка), рівний одній часовій одиниці. Таким чином, автокореляція першого порядку використовує перші  $N - 1$  спостережень  $x_t$ ,  $t = 1, \dots, N - 1$ , та наступні  $N - 1$  спостережень  $x_t$ ,  $t = 2, \dots, N$ . Кореляція між  $x_t$  та  $x_{t+1}$  визначається наступним чином:



$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - \overline{x_{(1)}})(x_{t+1} - \overline{x_{(2)}})}{\sqrt{\sum_{t=1}^{N-1} (x_t - \overline{x_{(1)}})^2} \sqrt{\sum_{t=1}^{N-1} (x_{t+1} - \overline{x_{(2)}})^2}}, \quad (2.5)$$

де  $\overline{x_{(1)}}$  є середнім для першого проміжку спостережень (перші  $N-1$  значень), а  $\overline{x_{(2)}}$  є середнім для другого проміжку (другі  $N-1$  значень). Коефіцієнт, що визначається у (2.4), і називається коефіцієнтом автокореляції.

Для суттєво великої кількості спостережень (великого  $N$ ) різниця між середнім на першому інтервалі  $\overline{x_{(1)}}$  та середнім на другому інтервалі  $\overline{x_{(2)}}$  є несуттєвою і може бути проігнорована, таким чином  $r_1$  буде приблизно рівним:

$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2}, \quad (2.6)$$

де  $\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t$  є середнім для досліджуваного періоду.

Рівняння (2.7) може бути узагальнене для отримання кореляції між спостереженнями, розділеними  $k$  часовими інтервалами:

$$r_k = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2}. \quad (2.7)$$

Значення  $r_k$  називається *коефіцієнтом автокореляції з лагом  $k$* . Графік функції автокореляції як залежності  $r_k$  від  $k$  називають *корелограмою*.

## 2.2. Хід виконання роботи

Після завантаження файлу з моделлю та виконанням всіх комірок отримуємо вікно для введення початкових даних (рис. 2.1).

## Аналіз флуктуацій часового ряду

hide

Оберіть дію

Завантажити дані Розрахувати прибутковість Взяти модуль ряду Нормалізувати

Волатильність Автокореляція Гаусовський розподіл Розподіл ймовірностей

Лівий хвіст Правий хвіст

Фонд

Кількість значень

Ширина вікна

Кількість інтервалів

Побудувати графік

Рис. 2.1. Інтерфейс моделі для проведення аналізу флуктуацій часового ряду

Оскільки завантаження даних відбувається з сайту [www.finance.com.ua](http://www.finance.com.ua), то для початку роботи необхідно у поле «Фонд» увести індекс фонду або акції, обрати «Кількість значень» та натиснути «Завантажити дані». У вікні з'явиться інформація про часовий ряд: кількість значень, мінімальне та максимальне значення, та інші основні статистики. У даній роботі використовуватимемо в якості зразка індекс Доу Джонса фондового ринку США за період, що містить кризу 1929 року (так звана «Велика депресія») (dj\_29). Часовий ряд має довжину 2000 днів, причому початок кризи приходить точно на середину ряду (точку 1000). Такий вибір дає можливість спостерігати характерні особливості поведінки похідних характеристик часового ряду у передкризові та післякризові періоди. Якщо вони носять універсальний характер, то ми можемо сподіватися на побудову передвісників або післявісників фінансових криз.

Використавши пункт «Побудувати графік», отримаємо рис. 2.2.

З рисунку видно, що ряд нестационарний та викликає певні ускладнення для подальшого аналізу. Тому перейдемо до прибутковостей («Розрахувати прибутковість»), які вже є стаціонарними, а їх нормалізація стандартним відхиленням («Нормалізувати») дозволяє легко порівнювати їх як з іншими часовими послідовностями, так і з розподілом Гауса.

На рис. 2.3 відображено поведінку нормалізованих прибутковостей.

Зверніть увагу, що флуктуації нормалізованих прибутковостей досить часто перевищують величину  $\pm 3$  (виділена суцільними лініями на рис. 2.3), що, як відомо, надзвичайно рідко спостерігається для незалежних подій. Цей факт можна відобразити шляхом порівняння функції розподілу нормалізованих флуктуацій з розподілом Гауса (рис. 2.4). Очевидно, що хвости розподілу вихідного ряду містять значні флуктуації, вони досить помітні (часто кажуть «важкі» у порівнянні з самою «головою» розподілу).

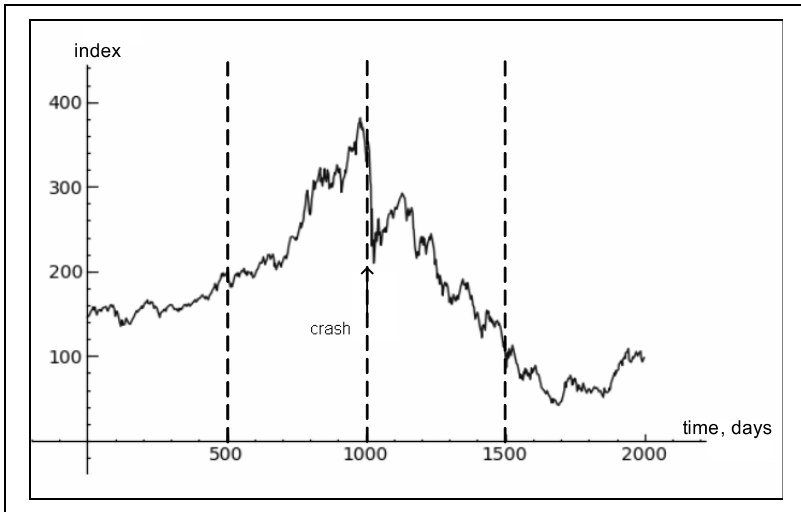


Рис. 2.2. Фінансовий часовий ряд індексу Доу Джонса 1927-1931 рр, який характеризується кризою (crash) у точці 1000.

Більш того, ми спостерігаємо відхилення флуктуацій від гаусівського розподілу саму у період кризи. Таке явище носить назву кластеризації волатильності, що буде виявлене пізніше і при дослідженні волатильності.

Отже, можна стверджувати, що прибутковості не є незалежними. Підтвердження цьому факту будемо шукати шляхом дослідження кореляційних властивостей обраного часового ряду.

Побудуємо функцію автокореляції. Для цього натиснемо кнопку «Автокорреляція», введемо початковий лаг (початкову часову затримку, «Початковий лаг»), крок лагу («Крок лагу») та кінцевий лаг («Кінцевий лаг»). Виберемо початковий, кінцевий лаг та крок лагу відповідно 0, 250, 1 (такі значення встановлено за замовчуванням) і

натиснемо кнопку «Побудувати графік». З'явиться графічне зображення корелограми.

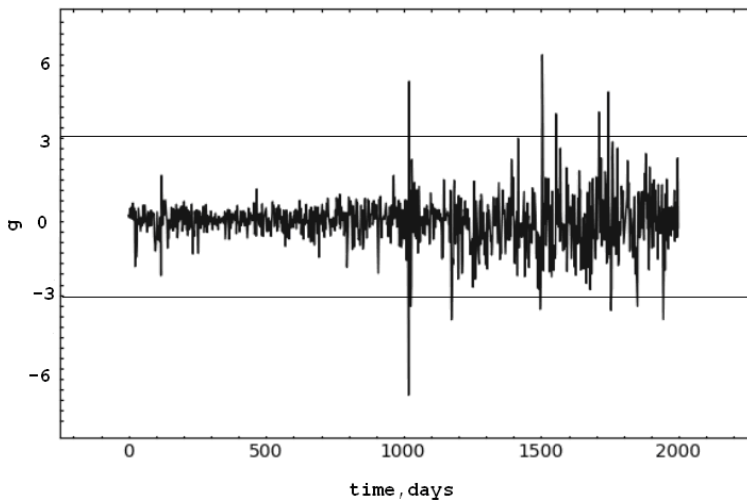


Рис. 2.3. Нормалізовані прибутковості обраного часового ряду

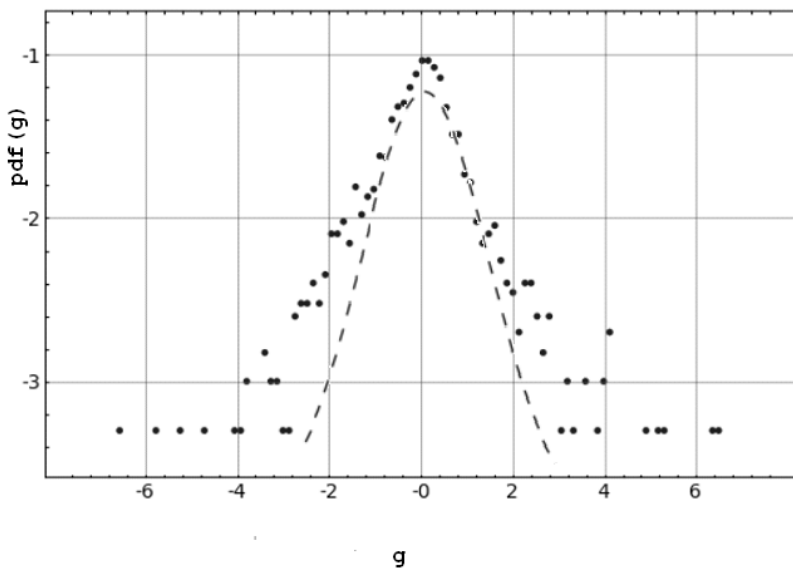


Рис. 2.4. Порівняння функцій розподілу нормалізованих прибу-

тковостей  $g$  з нормальним розподілом (штрихова лінія)

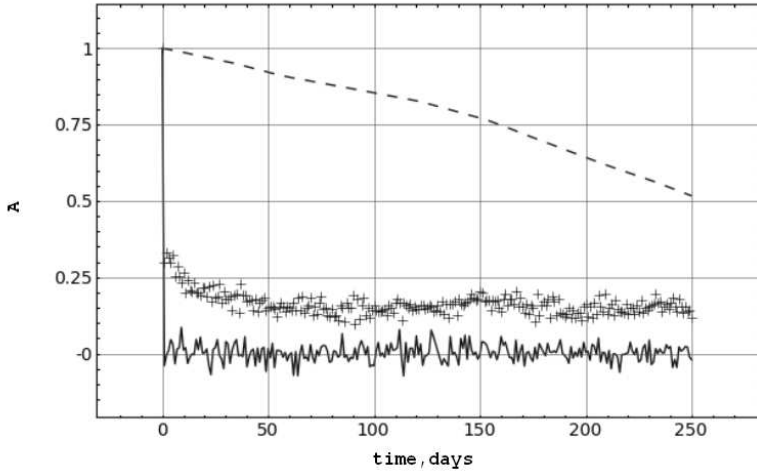


Рис. 2.5. Зміна з часом парних автокореляційних функцій для вихідного ряду (штрихова лінія), нормалізованих прибутковостей (суцільна) та модулів нормалізованих прибутковостей («+»)

Більш суттєву інформацію дають модулі прибутків, оскільки в даному випадку вимірюється лише наявність відхилень, а не їх характер (у який бік відхилення – в додатній чи від’ємний). На побудованому графіку корелограми можна добре бачити наявність інтервалу з додатними значеннями автокореляції. Зауважимо, що природа довготривалої пам’яті на сьогодні до кінця не досліджена, але очевидно, що для різних цінкових активів і їх похідних вона різна. Це треба враховувати в задачах прогнозування цінкових коливань.

З рис. 2.5 бачимо, що найбільш корельованими є ряд вихідних значень часового ряду. Навпаки, нормалізовані прибутковості мають найменшу пам’ять. Проміжна ситуація спостерігається у випадку модулів прибутковостей. Це свідчить про різну ступінь стаціонарності відповідних часових рядів.

Наступним кроком у дослідженні ряду є аналіз волатильності.

Згідно з (2.2) волатильність розраховується для модулів прибутковостей, тому виберемо послідовно у основному вікні «Завантажити дані», «Розрахувати прибутковість», «Взяти модуль ряду» і розрахуємо волатильність, натиснувши «Волатильність». Обираємо параметри «Ширина вікна» (ширина вікна для розрахунку середніх значень волатильності) та «Крок вікна». Для встановлених за

замовчуванням параметрів отримаємо результат відображений на рис. 2.6.

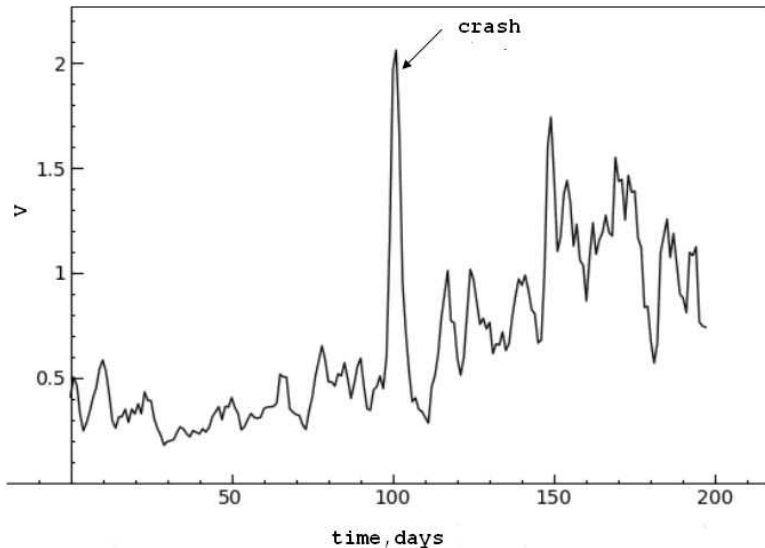


Рис. 2.6. Розрахована волатильність для модулів прибутковостей

Аналізуючи графік, можна зробити висновок, що у певні моменти спостерігалися стрибки волатильності із поступовим зменшенням її до попереднього рівня, що може бути наслідком збурень у процесі роботи ринку. Аналіз таких збурень, їх частоти та сили, дозволяє виявити приховані закономірності роботи ринку. У даному випадку максимальна волатильність спостерігається у кризовий період. Саме тому говорять про підвищену волатильність ринку в умовах кризи.

Таким чином, аналіз флуктуацій часового ряду, стаціонарного ряду прибутковостей, автокореляцій та волатильностей дозволяє отримати певні висновки, що можуть допомогти в роботі із часовими рядами, що досліджуються і ринком, з якого взято зазначені часові ряди. Зокрема, у даному випадку, можна давати рекомендації щодо поведінки фінансового ринку у кризовий період.

### 2.3. Завдання для самостійної роботи

1. Оберіть за допомогою викладача базу даних.
2. Проведіть дослідження часових флуктуацій, волатильностей

та автокореляцій згідно інструкції.

3. Проведіть дослідження автокореляцій прибутковостей для вихідних і перемішаних рядів.

4. Зробіть висновки.

5. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

#### **2.4. Контрольні питання**

1. Порівняйте вид залежностей флуктуацій цін і прибутковостей. Чому при розрахунках користуються не цінами, а прибутковостями?

2. Яку характеристику ряду визначає волатильність?

3. Знайдіть функцію розподілу значень волатильностей та порівняйте його із нормальним розподілом. Зробіть висновки.

4. У чому причина різних залежностей функцій автокореляцій для прибутковостей та їх модулів?

### 3. Мультифрактальний аналіз флуктуацій часового ряду

Існує кілька альтернативних підходів до оцінки фрактальної структури для часового ряду:  $R/S$  – аналіз; метод, заснований на визначенні клітинної розмірності; стандартний аналіз флуктуацій; аналіз детрендованих флуктуацій (АДФ) другого порядку; мультифрактальний АДФ. Розглянемо більш детально ці методи.

#### 3.1. R/S-аналіз

Метод R/S-аналізу, розроблений Б. Мандельбротом та А. Уоллесом [8.4], базується на попередньо створеному методі гідрологічного аналізу Херста, і надає можливість обчислювати параметр самоподібності  $H$ , що вимірює інтенсивність довготривалих залежностей у часовому ряді. Коефіцієнт  $H$ , який називають коефіцієнтом Херста, містить мінімальні прогнози стосовно природи системи, що вивчається, і може класифікувати часові ряди. За допомогою цього показника розрізняють випадкові та не випадкові ряди; окрім того, він пов'язаний із фрактальною розмірністю [8.4], що, у свою чергу, характеризує ступінь згладженості графіка, побудованого на основі часового ряду. Методом R/S-аналізу також можна виявити максимальну довжину інтервалу (цикл), на якому значення зберігають інформацію про початкові дані системи (довготривала пам'ять).

Аналіз починається із побудови ряду логарифмічних прибутків,

$$G_i(t) \equiv \ln \frac{S_i(t + \Delta t)}{S_i(t)}, \text{ де } S_i(t) - \text{значення вихідного часового ряду}$$

у момент  $t$ ,  $\Delta t$  – часовий крок. Отримана послідовність  $G_i(t)$  розбивається на  $d$  підпослідовностей довжини  $n$ . Для кожної підпослідовності  $m = 1, \dots, d$ :

1) шукається середнє значення  $E_m$  та стандартне відхилення  $S_m$ ;

2) дані нормалізуються шляхом віднімання середнього значення послідовності  $X_{i,m} = G_{i,m} - E_m$ ,  $i = 1, \dots, n$ ;

3) створюється послідовність накопичень  $Y_{i,m} = \sum_{j=1}^i X_{j,m}$ ,  
 $i = 1, \dots, n$ ;

4) знаходиться розмах



$R_m = \max\{Y_{1,m}, \dots, Y_{n,m}\} - \min\{Y_{1,m}, \dots, Y_{n,m}\}$ , який нормується середнім квадратичним відхиленням  $R_m/S_m$ ;

5) обчислюється середнє  $(R/S)_n$  нормованих значень розмаху для всіх підпослідовностей довжини  $n$ .

R/S-статистика, розрахована у такий спосіб, що відповідає співвідношенню  $(R/S)_n \cong cn^H$ , де значення  $H$  може бути отримане шляхом обчислення  $(R/S)_n$  для послідовності інтервалів зі збільшенням часового горизонту:

$$\log(R/S)_n = \log c + H \log n. \quad (3.1)$$

Знайти коефіцієнт Херста можна, побудувавши залежність  $(R/S)_n(n)$  у подвійному логарифмічному масштабі і взявши коефіцієнт нахилу прямої, що інтерполює точки отриманого графіка. Якщо значення  $H = 0,5$ , говорять про послідовність, що представляє собою білий шум;  $0,5 < H \leq 1$  свідчить про персистентний ряд, коли існує тенденція слідування великих значень ряду за великими і навпаки;  $H < 0,5$  вказує на антиперсистентний ряд.

При збільшенні часового горизонту коефіцієнт нахилу інтерполюючої прямої повинен прямувати до значення  $H = 0,5$ ; сам процес переходу свідчить про втрату впливу початкових умов на точні значення, і, таким чином, можна говорити про горизонт довгої пам'яті – це точка, до якої коефіцієнт нахилу інтерполюючої прямої відмінний від  $0,5$ , а після – близько  $0,5$ .

### 3.2. Стандартний аналіз флуктуацій

У стандартному аналізі флуктуацій (АФ) розглядається «профіль накопичень» (runoff profile)

$$y_n = \sum_{i=1}^n x_i, \quad n = \overline{1 \dots N} \quad (3.2)$$

і вивчається швидкість зміни флуктуацій профіля в залежності від зміни ширини вікна  $s$ . Можна розглядати профіль  $y_n$  як позицію точки у ланцюгу випадкових блукань після  $n$  кроків. Випадкові блукання починаються в деякій точці; на  $i$ -му кроці точка переміщується вправо, якщо  $x_i$  є додатнім, або ж вліво, якщо  $x_i$  – від'ємне.

Для знаходження зміни скейлінгового показника в залежності від  $s$  спочатку часовий ряд довжиною  $n$  елементів розбивається на під-

послідовності, що не перекриваються,  $N_s = \text{int}(N/s)$  довжиною  $s$ , починаючи з першого елемента, та  $N_s$  підпослідовностей, що не перекриваються, довжиною  $S$ , починаючи з останнього елемента. Потім визначаються флуктуації у кожному сегменті  $V$ .

У стандартному АФ флуктуації обчислюють безпосередньо від значень профілю на обох кінцях кожної підпослідовності  $V$ ,  $F^2(v, s) = (y_{vs} - y_{(v-1)s})^2$  і знаходиться середнє значення  $F^2(v, s)$  для  $2N_s$  підпослідовностей, щоб досягти реального значення флуктуації  $F_s(s)$ ,

$$F_2(s) \equiv \left( \frac{1}{2N_s} \sum_{v=1}^{2N_s} F^2(v, s) \right)^{1/2}. \quad (3.3)$$

За означенням,  $F_2(s)$  може розглядатись як середньоквадратичний зсув (переміщення) точки випадкових блукань в ланцюгу після  $s$  кроків. Для некорельованих значень  $x_i$  отримаємо закон дифузії А. Фіка  $F_2(s) \cong s^{1/2}$ . Для істотних випадків довгочасових кореляцій, де поведінка  $C(s)$  відповідає степеневому закону ( $C(s) = \langle x_i x_{i+s} \rangle \cong s^{-\gamma}$ ,  $0 < \gamma < 1$ ),  $F_2(s)$  збільшується теж згідно зі степеневим законом:

$$F_2(s) \cong s^H,$$

де флуктуаційний коефіцієнт  $H$  пов'язаний із кореляційним коефіцієнтом  $\gamma$  і коефіцієнтом спектра потужності  $\beta$  наступним законом:

$$H = 1 - \frac{\gamma}{2} = \frac{(1 + \beta)}{2}.$$

Для степеневих кореляцій, що спадають швидше за  $\frac{1}{s}$ , отримуємо  $H = \frac{1}{2}$  для великих значень  $s$ , так само як і для некорельованих даних.

Варто зауважити, що стандартний аналіз флуктуацій дещо подібний до R/S-аналізу, розробленого Херстом, за виключенням того,

що він базується на другому моменті  $F_2(s)$ , в той час як Херст розглядав випадок першого моменту  $F_1(s)$ . Для монофрактальних даних  $H$  ідентичний до коефіцієнта Херста.

### 3.3. Аналіз детрендованих флуктуацій

Аналіз детрендованих флуктуацій (АДФ) базується на гіпотезі про те, що корельований часовий ряд може бути відображений на самоподібний процес шляхом інтегрування. Таким чином, вимірювання властивостей самоподібності може непрямо свідчити про кореляційні властивості ряду. Переваги АДФ порівняно з іншими методами (спектральний аналіз, R/S-аналіз) полягають в тому, що він виявляє довгочасові кореляції нестационарних часових рядів, а також дозволяє ігнорувати очевидні випадкові кореляції, що є наслідком нестационарності.

Існують АДФ різних порядків, що відрізняються трендами, які вилучаються з даних.

Розглянемо АДФ найнижчого порядку.

Часовий ряд довжини  $N$  інтегрується,  $Y(k) = \sum_{i=1}^k (X_i - \bar{X})$ , де

$X_i$  –  $i$ -те значення часового ряду,  $\bar{X}$  – його середнє значення,  $k = 1, \dots, N$ . Ця дія повторює процес знаходження накопичень у R/S-аналізі. Отриманий ряд  $Y(k)$  розбивається на  $m$  підпоследовностей (вікон) однакової ширини  $n$  і для кожної підпоследовності (у кожному вікні) виконується наступне:

1) за допомогою методу найменших квадратів знаходиться локальний лінійний тренд  $Y_t(k)$ ;

2) підпоследовність детрендується шляхом віднімання значення локального тренду  $Y_t(k)$  від значень ряду  $Y(k)$ , що належать підпоследовності  $t$ ;

3) знаходиться середнє  $\bar{Y}_t$  детрендованих значень.

Для отриманих таким чином значень на всіх підпоследовностях знаходиться:

$$F_n = \sqrt{\frac{1}{m} \bar{Y}_t},$$

де  $n$  – кількість точок у підпоследовності (ширина вікна),  $m$  – кі-

лькість підпоследовностей,  $\overline{Y}_t$  – середнє детрендованих значень для підпоследовності  $t$ .

Вказана процедура повторюється для вікон різної ширини, внаслідок чого отримуємо набір пар точок  $(F_n; n)$ . Побудова залежності  $\log F(n)$  від  $\log n$  та інтерполяція отриманих значень прямою дає змогу обчислити показник скейлінга  $\alpha$ , що є коефіцієнтом кута нахилу інтерполяційної прямої і характеризує зміну кореляцій логарифмічних прибутків часового ряду при збільшенні часового інтервалу.

Порівняно із R/S-аналізом метод АДФ дає більші можливості інтерпретації скейлінгового показника  $\alpha$ : для випадкового ряду (перемішаного чи «сурогатного»)  $\alpha = 0.5$ ; при наявності лише короткочасових кореляцій  $\alpha$  може відрізнятись від 0.5, проте має тенденцію прямувати до 0.5 при збільшенні розміру вікна. Значення  $0.5 < \alpha \leq 1.0$  показує персистентні довгочасові кореляції, що відповідають степеневому закону;  $0 < \alpha < 0.5$  означає антиперсистентний ряд. Спеціальний випадок, коли  $\alpha = 1$ , означає наявність шуму  $1/f$ . Для випадків, коли  $\alpha \geq 1$ , кореляції існують, проте перестають відображувати степеневу залежність; випадок  $\alpha = 1.5$  свідчить про Броунівський шум, який можна отримати інтегруванням білого шуму.

У АДФ другого порядку (АДФ2) обчислюються відхилення  $F^2(v, s)$  профілю від графіка інтерполяційного поліному другого порядку. Таким чином, вилучаються впливи можливих лінійних та параболічних трендів для масштабів, більших за розглядувані. Взагалі, у АДФ порядку  $n$  обчислюються відхилення профілю від інтерполяційного поліному  $n$ -го порядку, що вилучає вплив всіх можливих трендів порядків до  $(n-1)$  для масштабів, більших від розміру вікна.

Потім обчислюється найближчий поліном  $y_v(s)$  для профілю на кожному із  $2N_s$  сегментів  $V$  і визначається відхилення:

$$F^2(v, s) \equiv \frac{1}{s} \sum_{i=1}^s (x_{(v-1)s+i} - y_v(i))^2. \quad (3.4)$$

Далі застосовується (3.3) для визначення середньої флуктуації  $F_2(s)$ .

Так як АФ та різні види АДФ мають різні детрендовані власти-

вості, їх композиція може використовуватись як для виявлення довгочасових кореляцій, так і для знаходження типів наявних у ряді трендів, що не визначається за допомогою традиційних методів, таких як спектральний аналіз.

### 3.3.1. Локальний АДФ

Метод локального АДФ виник з припущення про існування в часовій послідовності як персистентних, так і антиперсистентних проміжків.

Для отримання даних на основі локального АДФ використовується вікно постійного розміру, що залежить від максимального значення  $S$  у стандартному АДФ, коли ще було отримано стабільне значення  $\alpha$ . Розмір вікна повинен бути достатнім для використання стандартного методу АДФ, за допомогою якого і визначається локальний коефіцієнт  $\alpha$ . Після цього вікно пересувається по послідовності на певний (невеликий) крок (наприклад, 20 днів) і обчислення повторюються.

Аналіз динаміки локального коефіцієнта Херста може допомогти при інтерпретації суттєвих змін у динаміці системи.

### 3.3.2. Мультифрактальний АДФ

Стандартний АДФ використовується для визначення (моно-)фрактальних скейлінгових властивостей і довгочасових кореляцій в зашумлених нестационарних часових рядах. Проте багато економічних (біологічних, медичних тощо) об'єктів не демонструють простої монофрактальної скейлінгової поведінки, що може бути визначена одним коефіцієнтом. В деяких випадках існує кросовер («*crossover*») на часових шкалах  $s_x$ , що відділяє моделі з різною поведінкою, наприклад, довгочасові кореляції на малих масштабах часу  $s \ll s_x$ , та кореляції іншого виду чи некорельовану поведінку на більших масштабах  $s \gg s_x$ . В деяких випадках поведінка скейлінга ще більш складна, і існують різні значення коефіцієнтів скейлінга для різних частин послідовності (наприклад, для першої та другої половини послідовності). Трапляються ще більш складні випадки, коли розглядаються накладені одна на одну множини із різною фрактальністю. Для таких випадків необхідно обчислювати множину коефіцієнтів скейлінга для повного опису поведінки об'єкта. Тоді застосовується мультифрактальний аналіз детрендованих флуктуацій.

У загальному випадку процедура мультифрактального АДФ

(МФ-АДФ) складається із п'яти кроків. Перші три кроки, по суті, ідентичні процедури стандартного АДФ.

Нехай є послідовність  $x_k$  довжини  $N$ , що не має великої кількості нульових значень («compact support»).

*Крок 1.* Визначається профіль (накопичення):

$$Y(i) \equiv \sum_{k=1}^i (x_k - \bar{x}), \quad i = \overline{1..N}. \quad (3.5)$$

Віднімання середнього  $\bar{x}$  є необов'язковим, оскільки може бути виконане пізніше детрендуванням на третьому кроці.

*Крок 2.* Профіль  $Y(i)$  розбивається на  $N_s = \text{int}\left(\frac{N}{s}\right)$  сегментів (підпослідовностей) однакової довжини  $s$ , що не перекриваються. Оскільки загальна довжина послідовності  $N$  часто не ділиться націло на  $s$ , залишок в кінці послідовності, що є меншим за ширину вікна, відкидається. Для врахування відкинутої частини і використання, таким чином, усіх елементів послідовності, вищевказана процедура повторюється також, починаючи з кінця послідовності. Таким чином, разом буде отримано  $2N_s$  підпослідовностей.

*Крок 3.* Для кожної із  $2N_s$  підпослідовностей обчислюється локальний тренд методом найменших квадратів. Потім визначається відхилення

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s (Y((v-1)s+i) - y_v(i))^2 \quad (3.6)$$

для кожного сегмента  $V$ ,  $v = \overline{1..N}$  і

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s (Y(N - (v - N_s)s + i) - y_v(i))^2 \quad (3.7)$$

для кожного  $v = \overline{N_s + 1..2N_s}$ . Тут  $y_v(i)$  є інтерполюючий поліном на сегменті  $V$ . Для інтерполяції використовуються лінійні, квадратичні, кубічні поліноми чи поліноми вищого порядку (традиційно називаються АДФ1, АДФ2, АДФ3 і т.д.). Оскільки детрендування часового ряду відбувається відніманням значень полінома від реальних значень ряду, АДФ різних порядків відповідно відрізняються у своїх можливостях по вилученню тренду в ряді.

У (МФ-)АДФ $m$  ((МФ-)АДФ  $m$ -го порядку) вилучаються тренди профілю порядку  $m$  (або відповідно порядку  $(m-1)$  для вихідного часового ряду). Таким чином, порівняння результатів роботи АДФ рі-

зних порядків використовується для отримання типу тренду у вихідному часовому ряді.

*Крок 4.* Знаходиться середнє по всіх підпоследовностях для отримання функції флуктуацій  $q$ -го порядку:

$$F_q(s) = \left( \frac{1}{2N_s} \sum_{v=1}^{2N_s} (F^2(s, v))^{q/2} \right)^{1/q}, \quad (3.8)$$

де, взагалі кажучи, значення змінної  $q$  може бути довільним, за виключенням нульового. Для  $q = 2$  отримуємо стандартний метод АДФ. При розгляді питання, як впливає часова шкала  $S$  при різних значеннях  $q$  на узагальнену залежність  $F_q(s)$  від  $q$ . Для цього необхідно повторити кроки 2-4 для різних часових масштабів  $S$ . Цілком зрозуміло, що  $F_q(s)$  буде збільшуватись із збільшенням  $q$ . Звичайно,  $F_q(s)$  залежить також від порядку методу АДФ  $m$ . Згідно з побудовою,  $F_q(s)$  визначене лише для значень  $s \geq m + 2$ .

*Крок 5.* Визначається скейлінгова поведінка функції флуктуацій шляхом аналізу у подвійному логарифмічному масштабі залежності  $F_q(s)$  від  $q$ . Якщо послідовність  $x_i$  має довгочасові кореляції,  $F_q(s)$  збільшується із збільшенням  $s$  згідно степеневому закону:

$$F_q(s) \cong s^{h(q)}. \quad (3.9)$$

Взагалі, коефіцієнт  $h(q)$  повинен залежати від  $q$ . Для стаціонарних часових рядів,  $h(2)$  ідентичний коефіцієнту Херста. Таким чином, функцію  $h(q)$  можна назвати узагальненим коефіцієнтом Херста.

Для монофрактальної часової послідовності  $h(q)$  залежить від  $q$ , таким чином скейлінгова поведінка відхилень  $F^2(v, s)$  однакова на всіх сегментах  $V$  і процедура усереднення (3.8) дасть однакові значення скейлінгового коефіцієнта для всіх сегментів послідовності. Лише у випадку, коли масштаби малих і великих флуктуацій відрізняються, буде помітною залежність  $h(q)$  від  $q$ : якщо розглядати додатні значення  $q$ , сегменти  $V$  з великими значеннями  $F_s^2(v)$  (наприклад, великі відхилення від відповідних інтерполяційних поліномів) будуть домінувати у середньому значенні  $F_q(s)$ . Таким чином, для додатних значень  $q$   $h(q)$  описує скейлінгову поведінку сегментів із великими флуктуаціями. Зазвичай великі флу-

ктуації характеризуються меншими скейлінговими коефіцієнтами  $h(q)$  для мультифрактальних рядів. Навпаки, для від’ємних значень  $q$  сегменти  $\mathcal{V}$  з малими відхиленнями  $F_s^2(\mathcal{V})$  будуть домінувати у середньому значенні  $F_q(s)$ . Таким чином, для від’ємних значень  $q$   $h(q)$  описує скейлінгову поведінку сегментів з малими флуктуаціями, що зазвичай характеризуються більшим скейлінговим коефіцієнтом.

Однак, метод МФ-АДФ може визначати лише додатні узагальнені коефіцієнти Херста  $h(q)$ , і стає неточним для сильно антикорельованих сигналів, коли  $h(q)$  наближається до нуля. В таких випадках використовується модифікований МФ-АДФ. Більш простим шляхом для аналізу подібних даних може служити інтегрування часового ряду перед виконанням процедури МФ-АДФ (знаходження накопичень). Звідси, просте знаходження суми у (3.5), що описує профіль початкових даних, замінюється подвійним знаходженням суми,

$$\tilde{Y}(i) = \sum_{k=1}^i (Y(k) - \bar{Y}). \quad (3.10)$$

Після цього виконується процедура МФ-АДФ, описана вище, внаслідок чого отримується узагальнена функція флуктуацій

$$\tilde{F}_q(s) \cong s^{\tilde{h}(q)} = s^{h(q)+1}. \quad (3.11)$$

Таким чином, скейлінгова поведінка може бути точно визначена навіть у випадку, коли  $h(q)$  близьке до нуля (проте більше за  $-1$ ) для деяких значень  $q$ . Зауважимо, що  $\tilde{F}_q(s)/s$  відповідає  $F_q(s)$  у формулі (3.9). Якщо на кожному кроці (3.10) не віднімається середнє значення, така сума, скоріше за все, дасть квадратичний тренд у профілі  $\tilde{Y}(i)$ . В цьому випадку необхідно використовувати, щонайменше, МФ-АДФ другого порядку для вилучення такого штучного тренду.

### 3.4. Хід виконання роботи

У середовищі Sage завантажимо робочий аркуш для дослідження часового ряду. Після завантаження необхідно виконати всі комірки.

Після запуску робочого аркушу та завантаженні даних (необхідно натиснути «Завантажити дані») у вікні з’явиться інформація про



відкритий ряд (рис. 3.1).

Найпростіше дослідження, яке можна провести для даного ряду – R/S-аналіз. R/S-аналіз проводиться не для самого ряду, а для прибутковостей («*return*»), тому розрахуємо їх, натиснувши кнопку «Розрахувати прибутковість», а потім виберемо «Показник Херста», після чого необхідно обрати поле «Коефіцієнт» та вибрати значення коефіцієнта за допомогою повзунка або ввести значення в поле «Ширина вікна» (рис. 3.2).

```
Центральний момент = 1 = 0.0
Центральний момент = 2 = 195534.872171
Центральний момент = 3 = 52777177.2239
Центральний момент = 4 = 105743061696.0
-----
Середні:
Арифметичне = 1417.81341579
Квадратичне відхилення = 442.27602104
Квадратичне = 1485.16994118
Модуль = 347.186349938
-----
Максимальне значення = 2538.916
Мінімальне значення = 582.971
Розмах= 1955.945
Відносний розмах= 1.37955035424
Відносне відхилення= 0.244874499049
Відносний коефіцієнт варіації =
0.311942330432
Коефіцієнт асиметрії = 0.610050766319
Ексцес = 2.76361986483
```

Рис. 3.1. Початкові статистичні дані для завантаженого ряду

Результат аналізу відображається у подвійному логарифмічному масштабі. Розрахунок точок для отримання значень коефіцієнта Херста проведемо з використанням послідовності ширини вікон, де наступне значення більше у кілька разів за попереднє (такий спосіб найкращий для відображення у вибраному нами масштабі); на рис. 3.2 коефіцієнт обраний 1.1, тобто перша ширина вікон може бути, наприклад, 100, наступна 110, далі: 121, 133, 146, 160 і т.д.

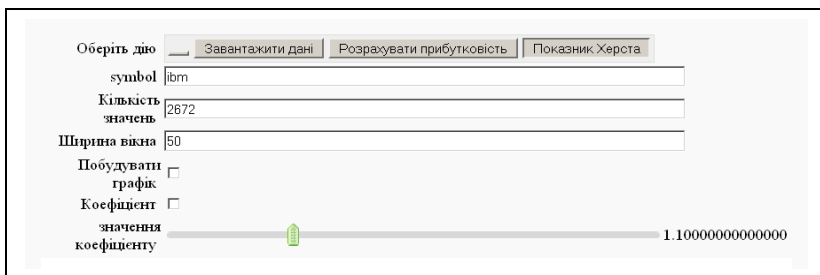


Рис. 3.2. Початковий вид програми з напівавтоматичним режимом керування для введення початкових даних

Після проведення розрахунків отримуємо графік, подібний до зображеного на рис. 3.3.

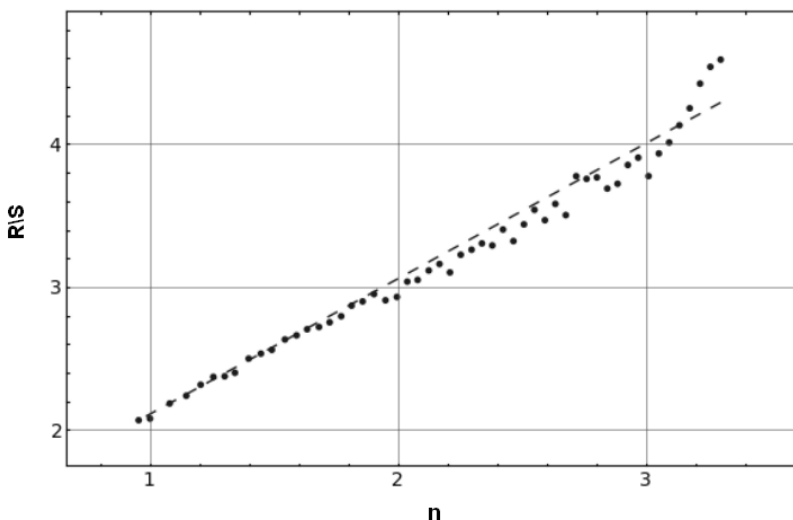


Рис. 3.3. Ряд, побудований на основі початкового, для визначення коефіцієнта Херста.  $H=0.71$

Значення коефіцієнта Херста відображається над графіком і рівний для нашого прикладу  $H=0.71 > 0.5$ , що свідчить про персистентність досліджуваного часового ряду.

Після проведення дослідження можемо закрити робочий аркуш і завантажити наступний, який дає більш важливу інформацію про ряд за допомогою аналізу методом детрендованих флуктуацій.

Для цього виберемо у основному вікні програми ряд «Розрахувати прибутковість» і знайдемо його модуль, вибравши «Взяти модуль ряду», для отриманого ряду «DFA» (Detrended Fluctuation Analysis). Дані, які необхідно ввести, дещо відрізняються від даних, необхідних для пошуку коефіцієнта Херста методом R/S-аналізу. Тут необхідно, окрім кроку збільшення вікна, вводити початковий та кінцевий лаг (часову затримку). Початковим значенням виберемо 5, а кінцевим 400 (можна було, звичайно, й більше, проте для зменшення часу обчислень встановимо саме таке значення). Для досягнення високої точності необхідно вибирати якомога більше значення, проте бажано не більше, ніж половина ряду. Після проведення вказаних дій натиснемо кнопку «DFA». Отримаємо зображення, подане на рис. 3.4.

Зверніть увагу на розходження значень коефіцієнта Херста, одержаних різними методами. Це говорить про різні можливості методів. Більш точним являється другий з них. Тому працюємо далі саме з цим методом.

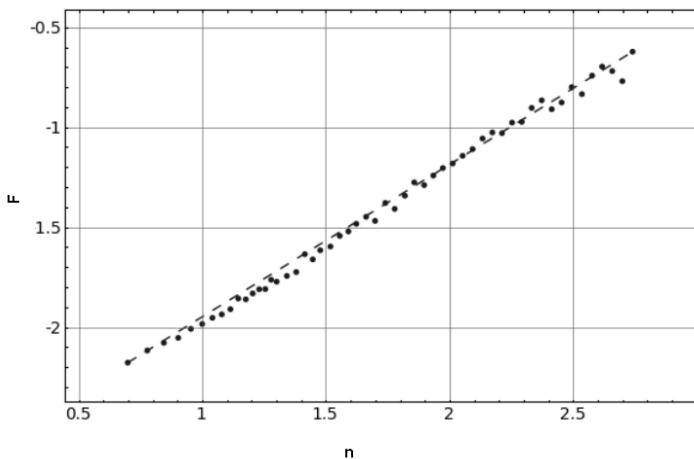


Рис. 3.4. Ряд, отриманий після обробки та знайдений коефіцієнт Херста за методом DFA ( $H=0,57$ )

Уважно подивившись на графік рис. 3.5, можна зробити висновок про наявність точки кросоверу (зміни кута нахилу прямої), наявність кросоверу свідчить про зміну акцентів у механізмах утворення довгої пам'яті; знаючи точку переходу, можна вже робити

припущення саме стосовно роботи ринку. Це також свідчить про наявність різних фрактальностей у часового ряду з різними коефіцієнтами Херста.

У загальному випадку говорять, що ряд характеризує мультифрактальну систему, яка характеризується не окремим значенням коефіцієнта Херста. Говорять про функцію розподілу коефіцієнтів  $H$ . Але перед тим як перейти до процедури мультифрактального аналізу, знайдемо так зване локальне значення коефіцієнта Херста. Для цього введемо поняття рухомого вікна певної довжини (наприклад, 250 днів, приблизно один рік), в якому розрахуємо коефіцієнт Херста (точка на графіку). Тепер перемістимо вікно вздовж ряду з кроком, наприклад, 1 день і знову знайдемо коефіцієнт Херста і так далі до кінця ряду. Оскільки часовий ряд постійно змінюється, характерну динаміку буде мати і значення  $H(t)$ . Для кризового ряду 1929 р. маємо результати, відображені на рис. 3.6.

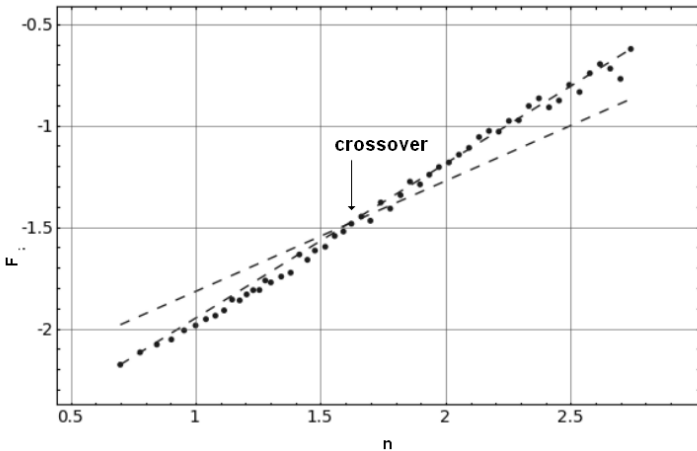


Рис. 3.5. Ряд, отриманий після обробки та знайдений коефіцієнт Херста за методом DFA для модулів прибутковостей

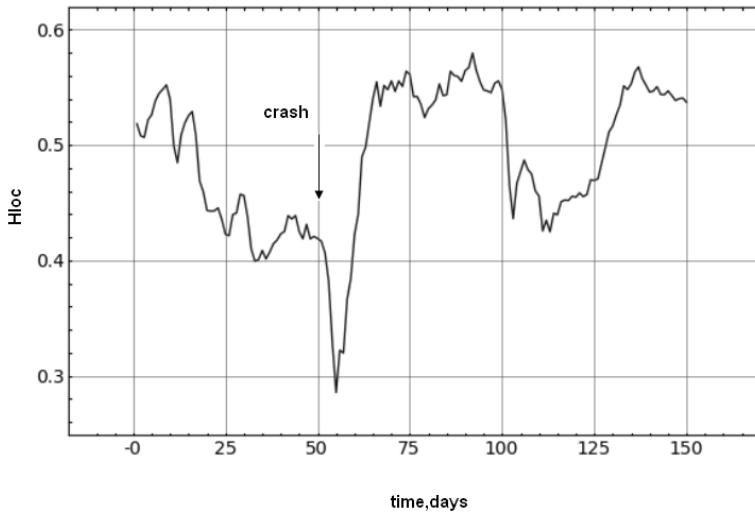


Рис. 3.6. Динаміка локального коефіцієнта Херста для часового ряду, який включає кризу 1929 року

У загальному випадку говорять, що ряд характеризує мультифрактальну систему, яка характеризується не окремим значенням коефіцієнта Херста. Говорять про функцію розподілу коефіцієнтів  $H$ . Але перед тим як перейти до процедури мультифрактального аналізу, знайдемо так зване локальне значення коефіцієнта Херста. Для цього введемо поняття рухомого вікна певної довжини (наприклад, 250 днів, приблизно один рік), в якому розрахуємо коефіцієнт Херста (точка на графіку). Тепер перемістимо вікно вздовж ряду з кроком, наприклад, 1 день і знову знайдемо коефіцієнт Херста і так далі до кінця ряду. Оскільки часовий ряд постійно змінюється, характерну динаміку буде мати і значення  $H(t)$ . Для кризового ряду 1929 р. маємо результати, відображені на рис. 3.6.

МФ-АДФ дозволяє визначити наявність мультискейлінга у ряді, що (у випадку існування) свідчить про використання для формування значень ряду композиції різних механізмів утворення цих значень.

Виберемо ряд для аналізу у завантаженому робочому аркуші. У цьому випадку відкриється вікно, в якому вибираються опції для проведення аналізу.

Кількість значень	10000
Початкова ширина сегмента	10
Кінцева ширина сегмента	1000
Степінь $q$ для розрахунку DFA	[-3,0,1,3]
Порядок DFA	2
Кількість вікон при кроці	100
Коефіцієнт збільшення вікна	1.1
Відсоток збільшення вікна	90
Тип збільшення вікна	<input type="radio"/> nSTEP <input type="radio"/> nRATIO <input type="radio"/> STEP <input type="radio"/> RATIO

Введіть всі поля даних

Рис. 3.7. Інтерфейс налаштування параметрів реалізації моделі МФ-АДФ

Для виконання МФ-АДФ у відкритому вікні необхідно провести відповідні налаштування. У полі «Початкова ширина сегмента» вказується найменша ширина сегментів  $s$ , на які розбивається вихідний ряд для проведення подальшого детрендування, а у полі «Кінцева ширина сегмента» – найбільша ширина сегментів. У полі «Степінь  $q$  для розрахунку DFA» вводиться найменше та найбільше значення степеня  $q$  та крок зміни значень з використанням синтаксису Sage для побудови значень ряду. Наприклад, рядок  $[-10,1,10]$  означатиме послідовність цілих чисел від  $-10$  до  $10$  із кроком  $1$ . Рекомендована послідовність степенів  $q = -3, -2.9, -2.8, \dots, 3$  ( $[-3,0.1,3]$ ). «Порядок DFA» вибирається у відповідному полі (рекомендується аналізувати за допомогою 2-го порядку). Необхідно обрати «Тип збільшення вікна», яке буде обирати можливі значення ширини сегменту  $S$  :

– якщо обрано «nRATIO», то у полі «Кількість вікон при кроці» встановлюється кількість різних можливих значень сегментів  $S$ , які вибиратимуться для проведення аналізу, причому буде розраховано коефіцієнт  $k$ , такий що,  $s_{i+1} = ks_i$ ;

– якщо обрано «STEP», то у полі «Крок збільшення вікна» встановлюється значення, на яке буде збільшено ширину сегмента наступної ітерації;

– якщо обрано «RATIO», то у полі «Коеф. збільшення вікна» встановлюється значення, на яке необхідно помножити  $S$  для отримання ширини наступних сегментів для подальшого видалення наступної ітерації;

– кнопка «nSTEP» аналогічна дії кнопки «nRATIO» за тією відмінністю, що в даному випадку крок обирається програмою таким

чином, щоб між шириною найменшого сегмента та найбільшого можна було утворити ще  $n - 2$  різних значень ширини (якщо це можливо), причому  $n$  береться із поля «Кількість вікон при кроці».

По закінченні розрахунку будується кілька функцій:

– у вікні «Розподіл  $F_q(s)$ » відображується розподіл залежності  $F_q$  від ширини сегменту  $S$  (рис. 3.8);

– у вікні «Розподіл  $H(q)$ » відображується розподіл залежності  $h(q)$  (рис. 3.9);

– у вікні «Розподіл  $\tau(q)$ » відображується розподіл залежності  $\tau(q)$  (рис. 3.10);

– у вікні «Розподіл  $F(\alpha)$ ,  $\alpha$  розрахована через  $\tau$ » відображується спектр мультифрактальності, отриманий за допомогою перетворення Лежандра (рис. 3.11).

Останні два пункти попереднього переліку (рис. 3.10 та рис. 3.11) і свідчать про мультифрактальність ряду.

Таким чином, для одного часового ряду можна отримати, використовуючи різні методи, сукупність показників, що характеризують досліджуваний об'єкт з точки зору впливу зовнішніх подій на систему, яка описується цим рядом.

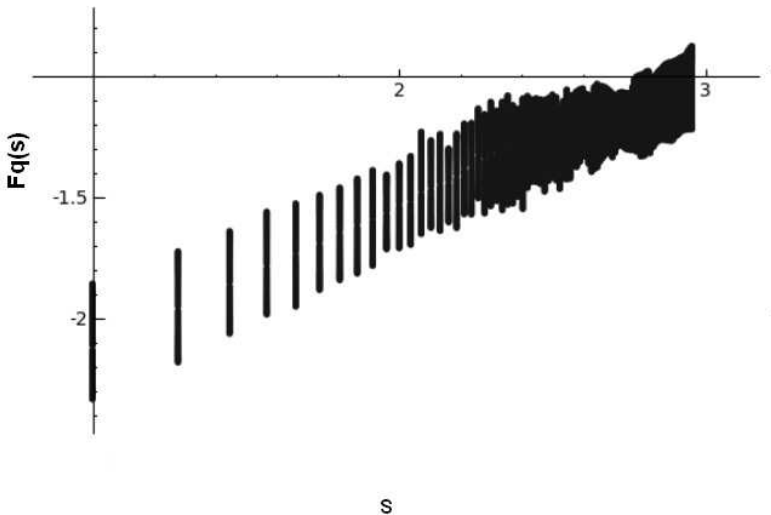


Рис. 3.8. Розподіл залежності  $F_q$  від  $S$

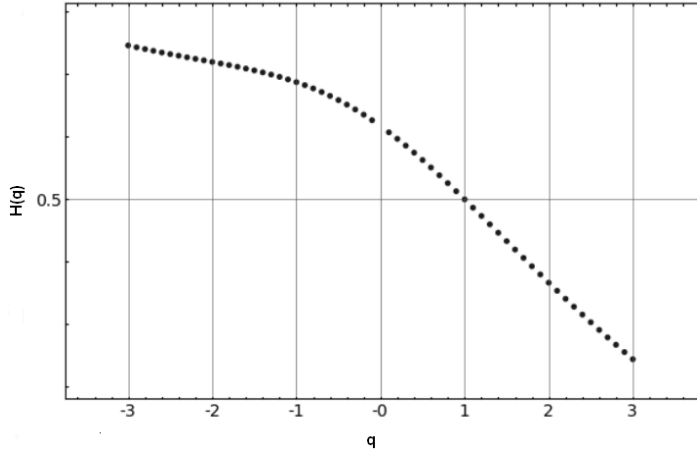


Рис. 3.9. Розподіл залежності  $H(q)$

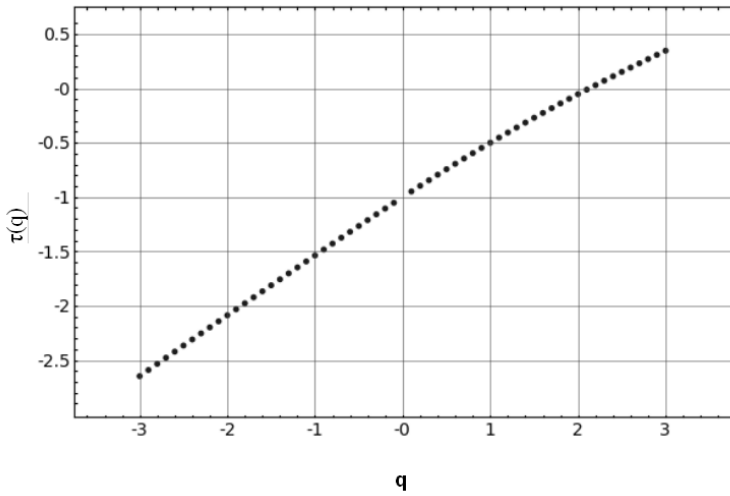


Рис. 3.10. Розподіл залежності  $\tau(q)$



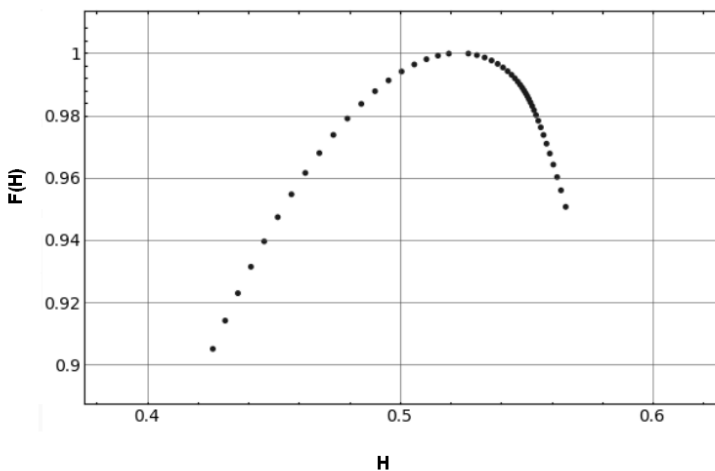


Рис. 3.11. Розподіл  $F(\alpha)$

### 3.5. Завдання для самостійної роботи

1. Виберіть із запропонованої бази даних варіант завдання.
2. Виконайте дослідження фрактальних характеристик заданих часових рядів.
3. Збережіть результати в окремому файлі.
4. Порівняйте значення коефіцієнтів Херста, одержаних методом R/S-аналізу та АДФ.
5. Проведіть повний аналіз ряду за допомогою методу мультифрактального АДФ.
6. Дайте інтерпретацію отриманим результатам.
7. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

### 3.6. Контрольні питання

1. У чому полягають переваги методу детрендованих флуктуацій перед методом R/S-аналізу?
3. Які ситуації досліджуваного ринку характеризує локальний коефіцієнт Херста?
2. На що вказує мультифрактальність часового ряду?

## 4. ДОСЛІДЖЕННЯ ПРОЦЕСІВ САМООРГАНІЗАЦІЇ В ЕКОНОМІЧНИХ СИСТЕМАХ

### 4.1. Моделювання колективних ефектів складних систем засобами теорії випадкових матриць

Вивчення статистичних властивостей матриць з незалежними випадковими елементами – *випадкових матриць* – має багату історію, що починається з ядерної фізики, де проблема з'явилася 60 років тому при вивченні енергетичних рівнів складних ядер, що існуючі на той час моделі були не в змозі пояснити. Теорія випадкових матриць (ТВМ) [8.4] була розвинена в цьому контексті Ю. Вігнером, Ф. Дайсоном, М. Мехтою та іншими для пояснення статистики рівнів енергії складних квантових систем. Дослідники постулювали, що функція Гамільтона, що описує важкі ядра, може бути задана матрицею  $H$  з незалежними випадковими елементами  $H_{ij}$ , отриманими з розподілу імовірності. Відштовхуючись від цього припущення, було зроблено низку вражаючих передбачень, що було підтверджено експериментально. Для складних квантових систем передбачення на основі ТВМ представляють середнє по всім можливим взаємодіям. Відхилення від універсальних передбачень ТВМ відображують системну специфіку, не випадкові властивості системи, забезпечуючи ключові підходи до розуміння базової взаємодії системи.

Недавні дослідження, що використовували методи ТВМ до аналізу властивостей матриці взаємних кореляцій  $C$ , показують, що близько 98% власних значень матриці  $C$  співпадають зі значеннями, що отримуються з використанням ТВМ, таким чином пропонуючи задовільний рівень хаотичності у вимірюваних крос-кореляціях. Також було знайдено, що існують відхилення від передбачень за допомогою ТВМ у близько 2% найбільших власних значень. Ці результати викликають наступні питання:

1. Яка можлива інтерпретація для відхилень від ТВМ?
2. Що можна сказати про структуру  $C$  з цих результатів?
3. Яке практичне значення отриманих результатів?

У даній роботі ці питання будуть вивчені докладно. Шляхом комп'ютерного моделювання виявлено, що найбільше власне значення матриці  $C$  представляє вплив усього ринку, що є звичайним для всіх акцій. Аналіз змісту власних значень, що відхиляються від ТВМ, показує існування взаємних кореляцій між акціями того ж самого

типу промисловості, найбільш капіталізованими акціями, і акціями фірм, що мають бізнес у певному географічному секторі (локалізовані територіально). Обчислюючи скалярний добуток власних векторів від одного періоду часу до наступного, можна побачити, що власні вектори, що відхиляються, мають різні ступені стабільності в часі, визначеному кількісно величиною скалярного добутку. Найбільші два-три власних вектори стійкі протягом тривалих періодів часу, у той час як для іншої частини власних векторів, що відхиляються, стабільність у часі зменшується як тільки відповідні власні значення наближаються до верхньої межі ТВМ [8.4].

#### 4.1.1. Знаходження коефіцієнтів матриці крос-кореляцій

Визначення кореляцій між різними акціями – тема, цікава не лише з точки зору наукових причин розуміння економіки як складної динамічної системи, але також і з практичних поглядів, зокрема, з точки зору розміщення активів і оцінки портфельного ризику [8.4]. Ми будемо аналізувати взаємні кореляції між акціями, застосовуючи поняття і методи теорії випадкових матриць, що використовуються в контексті складних квантових систем, де точний характер взаємодій між підодинамицями невідомий.

Для визначення кількісно кореляції спочатку обчислюється зміна цін («return», прибутковості) акції  $i = 1, \dots, N$  за час  $\Delta t$ ,

$$G_i(t) = \ln S_i(t + \Delta t) - \ln S_i(t), \quad (4.1)$$

де  $S_i(t)$  позначає ціну акції  $i$ . Оскільки різні акції мають різні рівні прибутковості (стандартні відхилення), визначатимемо нормалізоване повернення (нормалізовану прибутковість, «return»)

$$g_i(t) \equiv \frac{G_i(t) - \langle G_i \rangle}{\sigma_i}, \quad (4.2)$$

де  $\sigma_i \equiv \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2}$  – стандартне відхилення  $G_i$ , а  $\langle \dots \rangle$  позначає середнє значення за період часу, що досліджується. Тоді обчислення матриці взаємних кореляцій  $C$  зводиться до обчислення формули:

$$C_{ij} \equiv \langle g_i(t) g_j(t) \rangle. \quad (4.3)$$

Згідно з побудовою, елементи  $C_{ij}$  обмежені областю  $-1 \leq C_{ij} \leq 1$ , де  $C_{ij} = 1$  відповідає повним кореляціям,  $C_{ij} = -1$  – повним антикореляціям, і  $C_{ij} = 0$  свідчить про некорельованність пар акцій.

Труднощі в аналізі важливості та значення коефіцієнтів крос-кореляції  $C_{ij}$  виникають внаслідок кількох причин, що полягають в наступному:

– ринкові умови з часом змінюються і взаємна кореляція, що існує між будь-якою парою акцій, може бути непостійною (нестационарною);

– скінчена довжина досліджуваного ряду, доступного для оцінювання взаємних кореляцій, додає так званий «шум вимірювання»: чим коротше досліджуваний ряд, тим менш точними будуть отримані значення.

Якщо буде використано довгий ряд для розв'язання проблеми скінченої довжини, на отримані значення буде впливати нестационарність крос-кореляцій. З цих причин, емпірично виміряні крос-кореляції будуть містити «випадкові» складові, і найбільш важливою (і одночасно важкою проблемою) є оцінка в складі матриці  $C$  таких взаємних кореляцій, що не є результатом випадковості.

Яким же чином можна виділяти з  $C_{ij}$  ті акції, що залишилися корельованими на розглядуваному періоді часу? Щоб відповісти на це питання, перевіримо статистику  $C$  у порівнянні із так званою «нульовою гіпотезою» випадкової кореляційної матриці – матриці кореляцій, побудованої із взаємно некорельованих часових рядів. Якщо властивості  $C$  відповідають властивостям для випадкової матриці кореляцій, тоді можна говорити про те, що значення емпірично вимірюваних властивостей  $C$  випадкові. Навпаки, відхилення властивостей  $C$  від таких же властивостей для випадкової кореляційної матриці передає інформацію про «справжні» кореляції. Таким чином, нашою метою є порівняння властивостей  $C$  з такими ж властивостями випадкової матриці кореляцій і розділ властивостей  $C$  на дві групи: а) частина  $C$ , що відповідає властивостям випадкової кореляційної матриці («шум»), б) частина  $C$ , що відхиляється («інформація»).

#### 4.1.2. Розподіл власних значень

Для отримання інформації про взаємні кореляції  $C$  необхідно порівняти властивості  $C$  з такими ж властивостями випадкової матриці крос-кореляцій. У матричній нотації така матриця може бути виражена як

$$C = \frac{1}{L} GG^T, \quad (4.4)$$

де  $G$  – матриця розміру  $N \times L$  з елементами  $\{g_{im} = g_i(m\Delta t), i = 1, \dots, N; m = 0, \dots, L - 1\}$  і  $G^T$  позначає транспонування  $G$ . Розглянемо випадкову кореляційну матрицю

$$R = \frac{1}{L} AA^T, \quad (4.5)$$

де  $A$  – матриця розміру  $N \times L$ , що містить  $N$  часових рядів із  $L$  випадковими елементами  $a_{im}$  з нульовим середнім і одиничним відхиленням, що означають взаємну некорельованість.

Статистичні властивості випадкових матриць типу  $R$  відомі. Зокрема, у наближенні  $N \rightarrow \infty$ ,  $L \rightarrow \infty$ , такому, що  $Q \equiv \frac{L}{N} (> 1)$  фіксоване, показано аналітично, що функція розподілу щільності імовірності  $P_{rm}(\lambda)$  власних значень  $\lambda$  випадкової матриці кореляції  $R$  визначається як

$$P_{rm}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \quad (4.6)$$

для  $\lambda$  в межах границь  $\lambda_- \leq \lambda_i \leq \lambda_+$ , де  $\lambda_-$  і  $\lambda_+$  – найменше і найбільше власні значення  $R$ , відповідно,

$$\lambda_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}. \quad (4.7)$$

Звертаємо вашу увагу, що вираз (4.6) є точним для випадку розподілених за Гаусом матричних елементів  $a_{im}$ .

Порівняємо розподіл власних значень  $P(\lambda)$  для  $C$  з  $P_{rm}(\lambda)$ . Для цього обчислимо власні значення  $\lambda_i$  матриці  $C$ , причому  $\lambda_i$  впорядкуємо за зростанням ( $\lambda_{i+1} > \lambda_i$ ). При дослідженнях зверніть увагу на присутність чіткої «великої частини» власних значень, що спадають у межах границь  $[\lambda_-, \lambda_+]$  для  $P_{rm}(\lambda)$ . Також зверніть увагу на відхилення для деяких найбільших і найменших власних значень від отриманих за допомогою ТВМ.

Оскільки рівняння (4.6) є таким, що строго відповідає лише для  $L \rightarrow \infty$  і  $N \rightarrow \infty$ , необхідно перевірити також відхилення від ідеального випадку, оскільки робота проводиться завжди із скінченими рядами. При дослідженнях виявляється, що для кількох найбільших (найменших) власних значень ефект впливу скінчених величин  $L$  і  $N$  відсутній.

### 4.1.3. Обернене відношення участі

Вивчивши інтерпретацію найбільшого власного значення, що значно відхиляється від результатів ТВМ, зосередимось на власних значеннях, що залишаються. Відхилення розподілу компонентів власного вектора  $u^k$  від ТВМ Гаусового передбачення більш явне, коли відстань від верхньої границі ТВМ  $\lambda_k - \lambda_+$  збільшується. Оскільки близькість до  $\lambda_+$  збільшуються ефекти хаотичності, визначаємо кількість компонентів, що беруть значну участь в кожному власному векторі, що, у свою чергу, відображає ступінь відхилення від ТВМ для розподілу компонентів власного вектора. Для цього використовується поняття оберненого відношення участі (ОВУ), що часто застосовується в теорії локалізації. ОВУ власного вектора  $u^k$  визначається як

$$I^k = \sum_{l=1}^N [u_l^k]^4, \quad (4.8)$$

де  $u_l^k$ ,  $l = 1, \dots, N$  – компоненти власного вектора  $u^k$ . Значення  $I^k$  може бути проілюстровано двома граничними випадками:

- вектор з ідентичними компонентами  $u_l^k = \frac{1}{\sqrt{N}}$  має  $I^k = \frac{1}{N}$ ;
- вектор з одним компонентом  $u_1^k = 1$  і нульовими іншими має  $I^k = 1$ .

Таким чином, ОВУ визначає кількість даних з числа компонентів власного вектора, що значний впливають на ринок, заданий системою часових рядів.

Наявність векторів з великими значеннями  $I^k$  також виникає в теорії локалізації Андерсона. У контексті теорії локалізації часто знаходять «випадкову смугу матриць», що містять узагальнені стани з маленьким  $I^k$  в більшій частині спектра власних значень, тоді як основні стани локалізовані і мають великі  $I^k$ . Виявлення локалізованих станів для маленьких і великих власних значень матриці крос-кореляцій  $C$  нагадує про локалізацію Андерсона і припускає, що  $C$  може мати випадкову зону матричної структури.

## 4.2. Інтерпретація власних значень та власних векторів, що відхиляються від ТВМ

Кількість істотних внесків у власний вектор визначається (як уже

було описано) за допомогою ОВУ, для отримання ж загальних властивостей досліджуються компоненти  $\frac{1}{I^k}$  власного вектора  $u^k$ .

Пряма експертиза таких власних векторів, однак, не приводить до прямої інтерпретації їх економічної доречності. Щоб інтерпретувати їх значення, звернемо увагу, що найбільше власне значення значно більше, ніж інші  $N-1$  власні значення тоді як розмір матриці  $C$  рівний  $N$ . Таким чином, щоб аналізувати власні вектори, що відхиляються, необхідно видалити ефект найбільшого власного значення  $\lambda_{\max}$ .

Щоб позбутися ефекту  $\lambda_{\max}$ , і, відповідно,  $G^{\max}(t)$ , над прибутковостями кожної акції  $G_i(t)$  виконується регресія

$$G_i(t) = \alpha_i + \beta_i M(t) + \varepsilon_i(t), \quad (4.9)$$

або ж, у нашому випадку просте визначення тренда часового ряду, що відповідає  $\lambda_{\max}$  та видалення його з кожної часової послідовності. Тоді обчислюється матриця кореляції  $C$ , використовуючи формули (4.1)-(4.3). Далі обчислюються власні вектори  $u^k$ ,  $C$  і проводиться аналіз їх істотних учасників. Власний вектор  $u^{N-1}$  містить приблизно  $\frac{1}{I^{N-1}}$  істотних учасників, які є всіма акціями з великими значеннями ринкової капіталізації.

Далі проводиться аналіз істотних вкладників іншої частини власних векторів. Результатами аналізу у більшості випадків є те, що кожний з власних векторів, що відхиляються, містить акції, які належать подібним або пов'язаним галузям промисловості як істотним вкладникам. Ці власні вектори поділяють набір всіх акцій у різні групи, що містять акції з великою ринковою капіталізацією ( $u^{N-1}$ ), а також галузеві коаліції ( $u^{N-2}$ – $u^{N-9}$ ).

### 4.3. Хід виконання роботи

У середовищі Sage завантажимо робочий аркуш для дослідження кореляційних властивостей ринку, заданого часовими рядами його складових. У якості прикладу виберемо часові послідовності індексів інвестиційної привабливості країн розвинених та країн з економіками, які розвиваються (так звані індекси MSCI – Morgan Stanley Capital International). Виконання всіх комірок призводить до появи вікна програми (рис. 4.1).

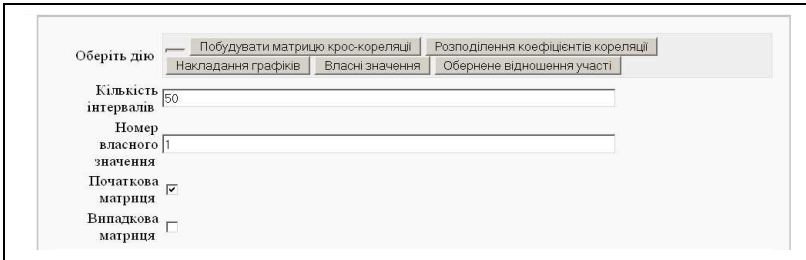


Рис. 4.1. Інтерфейс моделі для проведення дослідження ринків, заданих часовими рядами їх складових

Після завантаження файлу у вікні програми з'явиться інформація про відкриту базу, зокрема, ми можемо визначити, що відкрито матрицю з 48 рядками (кількість країн), та 2672 стовпчиками (кількість досліджуваних днів), та досить багато статистичної інформації про усі значення бази загалом (рис. 4.2).

Для проведення дослідження розрахуємо матрицю крос-кореляцій для чого обираємо «Побудувати матрицю крос-кореляції». Після побудови бачимо, що утворено матрицю розмірності  $N \times N$ , де  $N$  – кількість рядків, що знаходяться в матриці. Таким чином, знайдено матрицю  $C$  з використанням формул (4.1)-(4.3). Результат розрахунку зображено на рис. 4.3 у вигляді кольорової карти поля кореляцій.

```

Кількість строк 48
Кількість стовпців 2672
Центральний момент = 1 = 0.0
Центральний момент = 2 = 1710389.74492
Центральний момент = 3 = 5346764865.67
Центральний момент = 4 = 2.90166513005e+13
-----
Середні:
Арифметичне = 1024.26993177
Квадратичне відхилення = 1307.82379577
Квадратичне = 1661.17989334
Модуль = 920.139624715
-----
Максимальне значення = 10589.512
Мінімальне значення = 12.858
Розмах= 10576.654
Відносний розмах= 10.3260416731
Відносне відхилення= 0.898337045905
Відносний коефіцієнт варіації =
1.27683509513
Коефіцієнт асиметрії = 2.39025029468
Екссес = 9.91859941184

```

Рис. 4.2. Основні статистичні дані для відкритого ряду



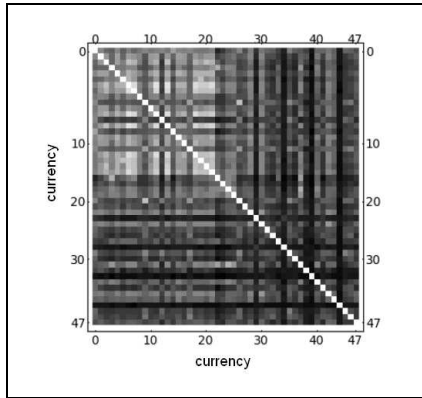


Рис. 4.3. Карта значень матриці крос-кореляції 48 індексів MSCI

Дослідимо статистичні властивості матриці  $C$  шляхом порівняння із властивостями випадкової матриці.

Знайдемо розподіл ймовірності елементів матриці  $C$ . Для цього введемо «Кількість інтервалів» рівним 50 та виберемо «Розподілення коефіцієнтів кореляції». У області графіка повинен з'явитись набір точок, що відповідає щільності ймовірності розподілу значень матриці крос-кореляцій (рис. 4.4).

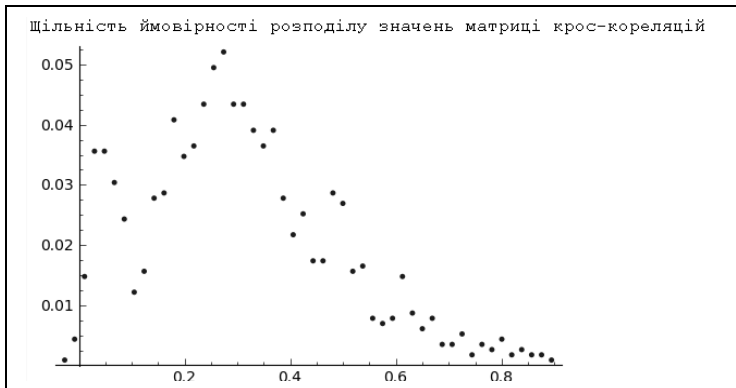


Рис. 4.4. Розподіл ймовірності значень початкової матриці крос-кореляцій

Для порівняння отриманого розподілу ймовірності з випадковою матрицею утворимо випадкову матрицю, перемішавши елементи

початкової матриці, у такий спосіб порушивши усі зв'язки (пам'ять), що існували(а) на ринку. Для цього зніmemo позначення з «Початкова матриця», а поставимо на поле «Випадкова матриця», після чого для випадкової матриці знайдемо так, як описано вище, матрицю крос-кореляцій та розподіл елементів цієї матриці (рис. 4.5–4.7).

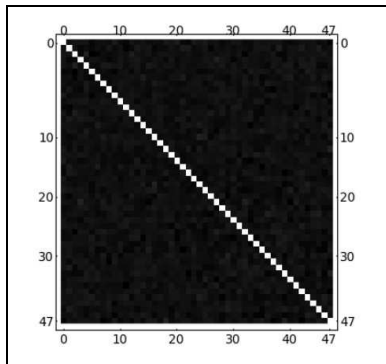


Рис. 4.5. Карта значень випадкової матриці крос-кореляції

Порівняємо розподіл ймовірності для початкової матриці та випадкової (рис. 4.6).

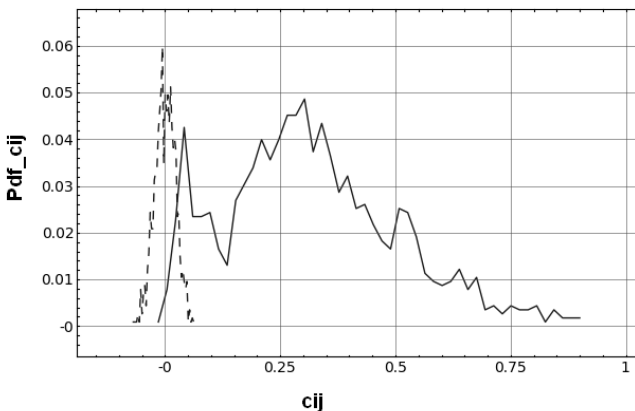


Рис. 4.6 Графік розподілу ймовірності для матриці крос-кореляції (пунктирна) та випадкової баз (суцільна)

Як видно із рис. 4.6, розподіл ймовірності значень матриці крос-кореляцій вихідного ряду зміщений далеко вправо порівняно з роз-

поділом для випадкової матриці, що свідчить про наявність кореляцій (а отже, і тісних зв'язків) між індексами окремих країн. Отже, шляхом досліджень властивостей матриці крос-кореляцій для початкової бази можна отримати статистично значущу інформацію про цей ринок та його складові.

Знайдемо розподіл власних значень матриці крос-кореляцій та проведемо роботу з ними.

Для знаходження власних найбільших значень та відповідних власних векторів виберемо матрицю крос-кореляцій початкової бази та пункт «Власні значення». У вікні з'явиться таблиця власних значень. Для вибору будь-якого власного значення в полі «Номер власного значення» обираєте номер. При виборі певного значення в полі «Номер власного значення» у полі справа відобразяться компоненти відповідного власного вектора.

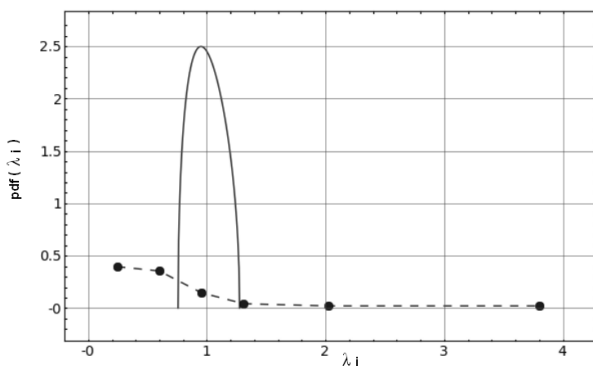


Рис. 4.7. Розподіл власних значень матриці взаємних кореляцій

На рис. 4.8–4.10 представлені три найбільші власні значення і відповідні власні вектори з їх компонентами.

Як можна побачити, існує найбільше власне значення, що суттєво більше за інші. Компоненти власного вектора вказують на країни, індекси яких найбільше впливають на інвестиційний ринок загалом. Крім того він характеризує групу країн Європи (європейський кластер).

Зверніть увагу, що компоненти власного вектора у верхній його частині (рис. 4.9) характеризують азійський кластер, тоді як компоненти власного вектора третього за величиною власного значення характеризують американський кластер (рис. 4.10).

Власне значення 17.1479758885

Номер	Власний вектор	Країна
1	0.211577096569	France
2	0.20449778217	UnitedKingdom
3	0.203620952778	Netherlands
4	0.202452134861	Italy
5	0.20161225734	Spain
6	0.195745178868	Germany
7	0.19288607003	Sweden
8	0.189108001804	Denmark
9	0.188707639799	Norway
10	0.188334723043	Switzerland
11	0.179693523942	Belgium
12	0.177064720915	Austria
13	0.176268000211	Portugal
14	0.167789984907	Ireland
15	0.164546535902	Hungary
16	0.161466583729	CzechRepublic
17	0.160479027647	Australia
18	0.158253049598	Finland

Рис. 4.8. Вікно для роботи із максимальним власним значенням

Власне значення 3.74171521618

Номер	Власний вектор	Країна
1	0.286665053225	HongKong
2	0.283152719248	China
3	0.268817990228	Korea
4	0.256268998681	Japan
5	0.254559132587	Singapore
6	0.237068780939	Taiwan
7	0.226910755213	Philippines
8	0.224655018226	Indonesia
9	0.223755719857	Malaysia
10	0.217109741108	Thailand
11	0.19806847911	Australia
12	0.177635855435	India
13	0.13334922535	NewZealand
14	0.128792434617	Egypt
15	0.0807031768193	Jordan
16	0.0574053997236	Russia
17	0.0537830132324	Pakistan

Рис. 4.9. Друге з найбільших власних значень та відповідні компоненти власного вектора

Власне значення 2.09842198593

Номер	Власний вектор	Країна
1	0.18486158121	Morocco
2	0.172506629098	Portugal
3	0.16112611287	Greece
4	0.151744209727	Austria
5	0.143320729103	Belgium
6	0.138616335481	Switzerland
7	0.132079929995	NewZealand
8	0.129388299164	Denmark
9	0.121328965925	Ireland
10	0.112872709899	Australia
11	0.110732250305	Italy
12	0.0982979616584	Spain
13	0.0924454002897	Netherlands
14	0.0918446409945	CzechRepublic
15	0.081482958606	France
16	0.0724894264886	Japan
17	0.0561474497472	Jordan

Рис. 4.10. Третє власне значення та відповідний власний вектор з його компонентами

Таким чином, за допомогою статистики власних значень та власних векторів можна робити висновок про взаємну кластерну організацію фінансового ринку та вплив загальної сукупності на окремі суб'єкти ринку і навпаки.

Знайдемо обернене відношення участі для матриці власних значень та векторів матриці крос-кореляцій початкової та матриці крос-кореляцій випадкової матриць.

Для обрахунку ОВУ початкової матриці натиснемо «Обернене відношення участі» (рис. 4.11).

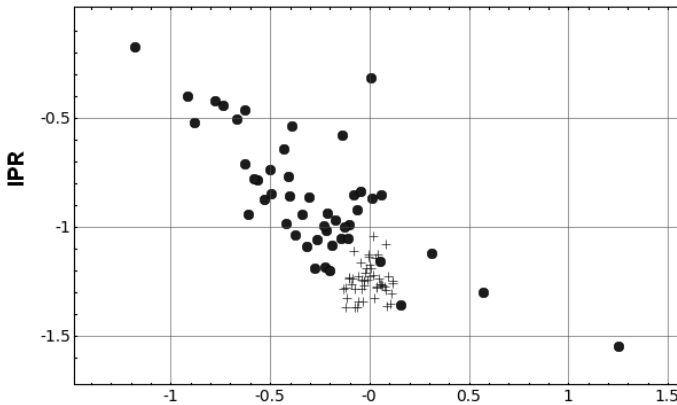


Рис. 4.11. Графіки ОВУ для вихідної (точки) та перемішаної («+») матриць

З рис. 4.11 добре видно, що графік ОВУ для перемішаної матриці добре локалізований, в той час як ОВУ для вихідної матриці дещо делокалізований на початку та наприкінці осі абсцис. Це свідчить про наявність інформації стосовно самоорганізації фінансових активів, їх взаємодії у сформованих кластерах.

Таким чином, при наявності сукупності часових рядів, що є даними діяльності економічних суб'єктів певного регіону (Європи, Азії тощо), можна провести дослідження стосовно структури вказаного регіону та взаємодії суб'єктів всередині неї. Дослідження проводяться на основі теорії випадкових матриць, що дозволяє отримувати інформацію шляхом аналізу матриці крос-кореляцій, побудованої для сукупної бази фінансово-економічних суб'єктів.

#### 4.4. Завдання для самостійної роботи

1. Виберіть матричну базу даних.
- 2 Проведіть дослідження за допомогою теорії випадкових матриць.
3. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

#### 4.5. Контрольні питання

1. Поясніть основну ідею теорії випадкових матриць.
2. Про що свідчить відмінність кореляційних і спектральних властивостей матриці даних і випадкової?
3. Дослідіть, як змінюється розподіл компонентів власного вектора у випадках: а)  $\lambda_- < \lambda < \lambda_+$ ; б)  $\lambda > \lambda_+$ ; в)  $\lambda < \lambda_-$ . Інтерпретуйте результати.
4. Порівняйте кольорову карту поля взаємних кореляцій випадкової матриці і заданої. Зробіть висновки.

## 5. КЛАСТЕРНИЙ АНАЛІЗ

### 5.1. Кластерний аналіз на основі часових рядів

У даній лабораторній роботі буде розглянуто, як саме визначити наявність економічної інформації стосовно взаємодії різних рівнозначних економічних об'єктів (ринків, акцій тощо), що міститься у часових рядах цін акцій.

#### 5.1.1. Гіпотеза ефективного ринку

У основній стохастичній моделі динаміки ціни акції приймається, що натуральний логарифм ціни є процесом, подібним до процесу дифузії. Емпіричні дослідження визначення степеню зміни кореляції під час розвитку системи показують, що часова кореляція слабка або взагалі відсутня на часових інтервалах, що варіюються від одного торгового дня до кількох років. Відсутність довготривалої пам'яті у функції парної автокореляції логарифмів зміни цін відповідає гіпотезі про ефективність економічного ринку. На ефективних ринках всі відомості (минулі, поточні та передбачення на майбутнє) доступні і негайно відбиваються у ціні товарів.

Моделювання флуктуацій цінкових часових рядів у термінах випадкових процесів може бути, на перший погляд парадоксальним, оскільки такий важливий індикатор, як ціна товару, є несуттєвим для випадкового процесу.

Розв'язання зазначеного парадоксу полягає в тому, що часовий ряд, багатий відомостями, є, насправді, незалежним від випадкових процесів. Таке твердження – один із висновків теорії складності. Внаслідок цього алгоритм дозволяє нам бути послідовними, одночасно заявляючи, що (1) часові цінні ряди є динамічними рядами, що містять досить великі обсяги даних і (2) вони добре моделюються в термінах випадкових процесів.

#### 5.1.2. Отримання даних з часової послідовності

Яким же чином можна отримати певні відомості з часової послідовності цін товарів?

Основною ідеєю є дослідження кількох часових послідовностей одночасно; тобто, дослідження зміни коефіцієнта кореляції з плином часу прибутковостей пар акцій  $i$  та  $j$ :

$$Y_i = \ln \frac{P_i(t)}{P_i(t-1)},$$

де  $i$  та  $j$  є порядковими номерами акцій, а  $P_i(t)$  є ціною закриття акції  $i$  в день  $t$ . Обчислення коефіцієнтів кореляції проводиться за наступною формулою:

$$\rho_{ij} = \frac{\langle Y_i Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{(\langle Y_i^2 \rangle - \langle Y_i \rangle^2)(\langle Y_j^2 \rangle - \langle Y_j \rangle^2)}}. \quad (5.1)$$

Кутові дужки  $\langle \dots \rangle$  – це середнє значення, розраховане для усіх днів розглядуваного часового періоду.  $\rho_{ij}$  визначає кореляцію, що спостерігається між логарифмом ціни акції  $i$  та логарифмом ціни акції  $j$ .  $\rho$  може змінюватись від -1 до 1.  $\rho = 1$  свідчить про існування повної кореляції між змінними  $Y_i$  та  $Y_j$ ,  $\rho = -1$  означає повну антикореляцію між  $Y_i$  та  $Y_j$ , а 0 спостерігається для некорельованих пар акцій.

Портфель із  $n$  акцій асоціюється із квадратною матрицею кореляцій розмірності  $n \times n$ , що є симетричною, з 1 на головній діагоналі. Обов'язкове число коефіцієнтів кореляції, що повністю описують ринок, рівне  $\frac{n(n-1)}{2}$ .

## 5.2. Побудова матриці відстаней на основі матриці крос-кореляцій

Для отримання певних відомостей з часових рядів, які можна інтерпретувати, необхідно спочатку визначити метрику, в межах якої проводитиметься інтерпретація. В якості функції від коефіцієнтів кореляції можна використати узагальнену відстань, експериментальним визначенням якої є

$$d(i, j) = 1 - \rho_{ij}^2. \quad (5.2)$$

Замість вказаної можна використовувати також дещо видозмінену формулу:

$$d(i, j) = \sqrt{2(1 - \rho_{ij})}. \quad (5.2')$$

У такому випадку в дослідженнях  $d(i, j)$  чисельно задовольняє трьом аксіомам Евклідової метрики:

$$d(i, j) = 0, \text{ якщо } i \text{ лише якщо } j \quad (*)$$



$$i = j;$$

$$d(i, j) = d(j, i); \quad (**)$$

$$d(i, j) \leq d(i, k) + d(k, j), \quad (***)$$

і тоді  $d(i, j)$  може бути використана в якості Евклідової відстані для всіх практичних задач. Перша аксіома виконується внаслідок того, що  $d(i, j) = 0$ , лише якщо кореляція повна ( $\rho = 1$ , що виникає у випадку, коли обидва економічних об'єкти виконують один і той самий стохастичний процес). Друга аксіома виконується, оскільки матриця коефіцієнтів крос-кореляцій і, відповідно, матриця відстаней  $D$  є симетричними за означенням. Третя аксіома виконується, оскільки вираз (5.2') еквівалентний Евклідовій відстані між двома векторами  $\widetilde{Y}_i$  та  $\widetilde{Y}_j$ , отриманими із часових рядів  $Y_i$  та  $Y_j$ , якщо використовувати кожний запис часового ряду як компонент вектора. Отриманий вектор мусить мати одиничну норму, а саме, він має бути отриманий відніманням від кожного запису середнього значення для всього ряду та нормалізацією його стандартним відхиленням.

Тепер, зважаючи на наведені міркування, можна дати практичне визначення метричної відстані між двома акціями певного портфеля. Однак метрична відстань є недостатньою для однозначного визначення таксономії, тому розглянемо більш обмежуючий топологічний простір, який буде однозначно асоційовано з таксономією. Таким обмежуючим топологічним простором є субдомінантний ультраметричний простір, асоційований з матрицею Евклідових відстаней. Ультраметрична відстань визначається трьома аксіомами. Перші дві співпадають з описаними у визначенні вище, тоді як нерівність трикутника, (\*\*\*) , замінюється більш сильною нерівністю, названою ультраметричною нерівністю:

$$d(i, j) \leq \text{Max}\{d(i, k), d(k, j)\}. \quad (5.3)$$

В просторі елементів, де визначена метрика, завжди можна отримати субдомінантну ультраметрику. Приклад процедури отримання матриці субдомінантних ультраметричних відстаней з матриці Евклідових відстаней можна знайти у [8.4]. В ультраметричному просторі елементи розміщуються не вздовж однієї лінії, а в ієрархічному

дереві. Набір акцій розглядуваного портфеля може інтерпретуватись як абстрактний простір,  $n$  елементів якого можуть бути з'єднані графом. У з'єднаному графі із взятими відстанями між усіма точками абстрактного простору мінімальне остівне дерево – МОД (minimal spanning tree – MST) має  $n-1$  ребро і найменшу довжину із довжин всіх остівних дерев, виходячи з суми відстаней між двома точками. Структура ієрархічного дерева отримаємо з матриці ультраметричних відстаней та МОД, асоційованого з нею.

Мінімальне остівне дерево та асоційоване з ним ієрархічне дерево показують існування кластерів будь-яких активів ринків, відомості про які мають велике значення з економічної точки зору. Отримана таксономія надає можливість групувати економічні об'єкти, що є гомогенними у відношенні до економічної активності.

### 5.3. Побудова мінімального остівного та ієрархічного дерев

Розглянемо систему індексів MSCI яка включає 49 країн, з яких 23 є розвиненими ([www.msci.com](http://www.msci.com)). Для побудови мінімального остівного дерева, асоційованого з Евклідовою відстанню, необхідно виконати наступні дії.

Спочатку усі елементи матриці відстаней  $D$  розташовуються у порядку неспадання (кожний наступний більший або рівний попередньому). Наприклад, маємо наступні 8 відстаней:

Пара країн	$d$
Іспанія-Італія	$d=0,61177$
Франція-Італія	$d=0,64137$
Голландія-Франція	$d=0,6462$
Германія-Франція	$d=0,65383$
Франція-Іспанія	$d=0,69175$
Аргентина-Бразилія	$d=0,73039$
Швеція-Італія	$d=0,77672$
Бельгія-Франція	$d=0,80568$
Мексика-Аргентина	$d=0,8217$

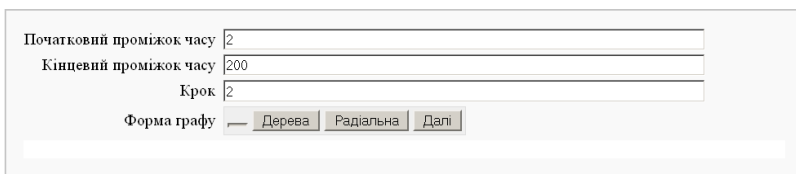
Мінімальне остівне дерево послідовно будується шляхом з'єднання усіх елементів у граф, виходячи із мінімальних відстаней між країнами на ринку.

Почнемо із пари елементів з найменшою відстанню. У даному випадку це дві країни: Іспанія та Італія ( $d=0,61177$ ). Наступною найменшою за порядком відстанню є відстань між елементами Франція та Італія ( $d=0,64137$ ), тому додаємо до МОД елемент Франція та з'єднуємо його з елементом Італія. Таким чином, отримане дерево вже буде мати вигляд Франція-Італія-Іспанія. Наступною парою елементів у відсортованому списку є країни Голландія-Франція, включаємо елемент Голландія до дерева. Аналогічно додаємо країну Германія. Наступною парою країн з найменшою відстанню є Франція-Іспанія, проте обидві країни вже включено в мінімальне остівне дерево, тому для цієї пари нічого не робимо. Аналогічно продовжуємо процедуру, поки всі елементи не будуть включені в дерево.

Таким чином, у результаті перебору усіх відстаней, отримаємо мінімальне остівне дерево, що буде характеризувати взятий ринок, стосовно того, які фірми взаємодіють між собою тісніше, на скільки груп якого характеру розділений ринок і т.д.

#### 5.4. Хід виконання роботи

У середовищі Sage завантажимо робочий аркуш для дослідження структури ринку, заданого часовими рядами його складових, виконавши всі комірки. Внаслідок цього відкриється вікно програми (рис. 5.1).



Початковий проміжок часу

Кінцевий проміжок часу

Крок

Форма графу  Дерева  Радіальна  Далі

Рис. 5.1. Інтерфейс моделі для проведення кластерного аналізу

На рис. 5.1 виділено три частини області користувача:

– кнопки вибору форми відображення мінімального остівного дерева системи;

– область, де задається часовий проміжок, за яким на екрані робляться всі обчислення.

Після завантаження даних необхідно вказати проміжок часу, за яким потрібно побудувати дерево. Перший день може бути не менший 2-го. Наприклад, цікавить проміжок від 2 до 204. Для цього

введіть наступне (рис. 5.2):

Початковий проміжок часу	<input type="text" value="2"/>
Кінцевий проміжок часу	<input type="text" value="200"/>
Крок	<input type="text" value="2"/>
Форма графу	<input type="radio"/> Дерева <input type="radio"/> Радіальна <input type="radio"/> Далі

Рис. 5.2. Інтерфейс моделі для проведення кластерного аналізу

Поля «Початковий проміжок часу» (початок) та «Кінцевий проміжок часу» (кінець) вказують вибраний проміжок часового ряду. Для переміщення необхідно збільшити пропорційно обидва поля. Спостерігайте за зміною виду дерев.

Після введення даних, на екрані з'явиться граф в обраній формі (рис. 5.3–5.4). При цьому різними кольорами зображені країни, які відносяться до різних регіонів. Це дозволяє виділити відповідні кластери і порівняти їх із знайденими раніше методом теорії випадкових матриць.

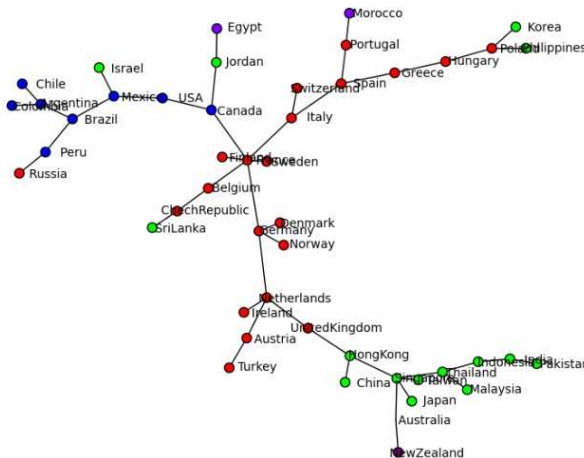


Рис. 5.3. Граф (мінімальне остівне дерево) початкової матриці MSCI

В даній роботі також є можливість досліджувати динаміку дерев, кластерів та їх характеристик у рамках алгоритму рухомого вікна. При дослідженні кризових явищ за допомогою теорії графів, ми бу-

демо переміщувати вибраний відрізок та спостерігати зміни зображень графів. У випадку, коли кінець вибраного проміжку дійде до дня кризи на ринку, який досліджується, помітно зміни у зв'язках між елементами системи.

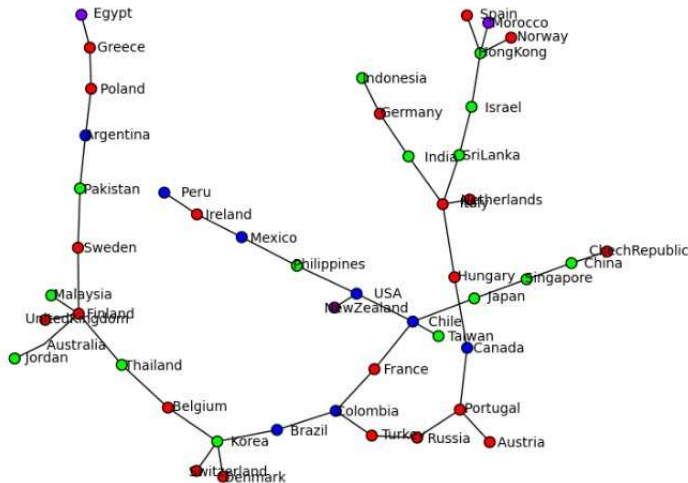


Рис. 5.4. Граф (мінімальне остівне дерево) для перемішаної матриці MSCI

У такий спосіб можна досліджувати структуру і динаміку складних систем, відслідковувати їх зміни з часом, аналізувати вплив критичних явищ.

### 5.5. Завдання для самостійної роботи

1. Виберіть відповідну базу даних, проведіть дослідження – кластерної структури, її динаміки, поясніть і інтерпретуйте результати;
  - порівняйте одержані результати з кластерною структурою, отриману методами теорії випадкових матриць
2. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

### 5.6. Контрольні питання

1. Поясніть основну ідею кластерного аналізу.
2. Про що свідчить формування відповідних кластерів ?

3. Як проявляються процеси глобалізації на кластерній структурі відповідних баз активів?

## 6. Дослідження динаміки складних систем за характеристиками кластерного дерева

### 6.1. Коефіцієнт кластеризації графа

Аналіз структури складної системи по її графам має суб'єктивний характер, оскільки важко визначити зміни в усій системі, розглядаючи її граф. Необхідно оперувати певними числовими характеристиками системи, по зміні яких можна робити висновок про стан системи. Такими характеристиками можуть виступати коефіцієнт кластеризації, діаметр графа, середня довжина шляху на графі. Розглянемо докладніше ці коефіцієнти та спосіб їх обчислення.

В суспільстві існують такі групи знайомих, коли кожен знає всіх інших в групі. Такі групи в теорії графів називають кластерами. Коефіцієнт кластеризації чисельно характеризує це явище в графі.

Існують різні способи визначення коефіцієнту кластеризації. Розглянемо деякі з них.

1. Візьмемо деяку вершину з графу та будемо обчислювати коефіцієнт кластеризації цієї вершини. До цієї вершини безпосередньо приєднані  $k_i$  вершин. Між цими вузлами існує  $E_i$  ребер. Максимально можлива кількість ребер, при якій ці вершини утворюють кластер, дорівнює  $k_i(k_i+1)/2$ .

**Коефіцієнт кластеризації вершини** є відношення між числом  $E_i$  ребер, що фактично існують між даними вершинами, до кількості ребер, при якому б ці вершини утворили кластер, дає величину коефіцієнта кластеризації вершини:

$$C_i = \frac{E_i}{E_{\max}} = \frac{2E_i}{k_i(k_i+1)}. \quad (6.1)$$

Розглянемо приклад графу (рис. 6.1).

Будемо обчислювати коефіцієнт кластеризації елемента France. Безпосередньо з ним з'єднано 5 країн. Між 6 країнами максимально можлива кількість зв'язків буде  $E_{\max}=6*5/2=15$ . Реально існує 8 зв'язків. Коефіцієнт кластеризації вузла France обчислюємо за формулою:

$$C_i = \frac{E_i}{E_{\max}} = \frac{2E_i}{k_i(k_i+1)} = \frac{2 \cdot 8}{6 \cdot 5} \approx 0,53.$$

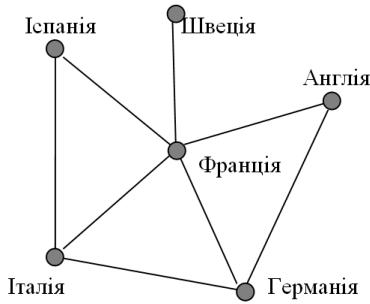


Рис. 6.1. Приклад зображення графу

Коефіцієнт кластеризації графу обчислюється як середнє арифметичне коефіцієнтів кластеризації всіх вершин.

2. **Коефіцієнт транзитивності.** Альтернативний до коефіцієнта кластеризації і обчислюється як відношення кількості трикутників (3 вершини, з'єднані одна з одною) до кількості з'єднаних пар ребер.

## 6.2. Середня довжина шляху на графі

Між кожною парою елементів системи на довільному графі можливі різні шляхи. Наприклад, розглянемо приклад з попередньої роботи. Нехай дано підрахунки відстаней між трьома елементами системи у вигляді матриці відстаней (табл. 6.1.)

Табл. 6.1

**Відстані між трьома елементами системи у вигляді матриці відстаней**

	<b>Іспанія</b>	<b>Італія</b>	<b>Франція</b>
<b>Іспанія</b>	0	0.31	0.69
<b>Італія</b>	0.31	0	0.34
<b>Франція</b>	0.69	0.34	0

Нам необхідно знайти матрицю мінімальних відстаней між кожними двома елементами. Зверніть увагу на те, що відстань від елемента Франція до елемента Іспанія дорівнює 0.69, а якщо розглянути шлях Франція-Італія-Іспанія, то відстань цього шляху буде дорівнювати сумі відстаней Франція-Італія та Італія-Іспанія. В нашому випадку це  $0,31+0,34=0,65$ , що є менше, ніж безпосередньо відстань Франція-Іспанія. З цих даних ми можемо утворити матрицю найкоротших (оптимальних) відстаней в графі.



**Матриця оптимальних відстаней в графі**

	<b>Іспанія</b>	<b>Італія</b>	<b>Франція</b>
<b>Іспанія</b>	0	0.31	0.69
<b>Італія</b>	0.31	0	0.34
<b>Франція</b>	0.69	0.34	0

Серед цих шляхів знаходиться середнє арифметичне, яке і буде наступним показником, за яким ми будемо слідкувати.

**6.3. Діаметр графа**

Діаметр графа – це найдовша відстань у графі. Її будемо знаходити з матриці оптимальних шляхів, приклад знаходження якої був вище розглянутий.

В нашому випадку діаметром графа буде відстань 0,65.

**6.4. Максимальна степiнь вершин графа**

Степенем вершини графа називається кількість ребер, які безпосередньо виходять із даної вершини. Розглянемо приклад дерева (рис. 6.2)

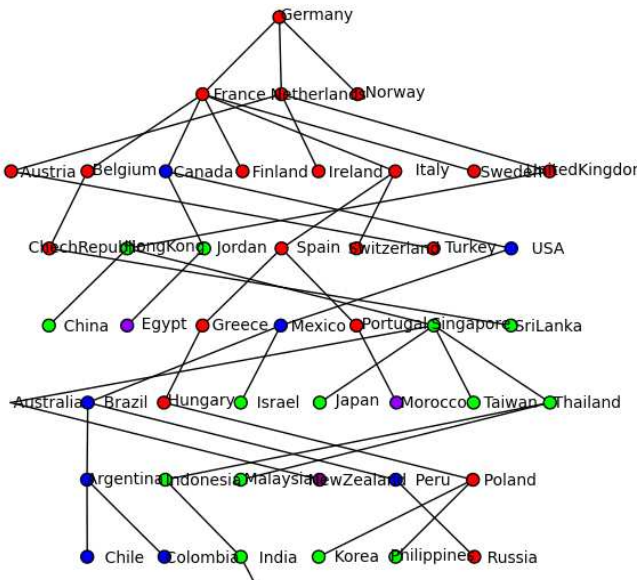


Рис. 6.2. Зображення графу у вигляді дерева

Найвищий на графі елемент – ХОМ має степінь 34, тому що безпосередньо з'єднаний з 34-ма фірмами. Елемент France має степінь 5 і т.д. Найвищий степінь у елемента France. Тому показником параметру «Найвища степінь у графі» буде число 5.

### 6.5. Хід виконання роботи

Аналогічно до попередніх робіт, у середовищі Sage завантажимо робочий аркуш для дослідження структури ринку, виконавши всі комірки, отримаємо вікно програми (рис. 6.3). Завантажуємо потрібну базу індексів та будуємо дерева на довільному проміжку.

The screenshot shows a software interface with the following elements:

- A button labeled "Оберіть дію" (Select action) with a dropdown menu currently showing "Розрахувати" (Calculate).
- A text input field for "Початковий проміжок часу" (Initial time interval) with the value "2".
- A text input field for "Кінцевий проміжок часу" (Final time interval) with the value "200".
- A text input field for "Довжина вікна" (Window length) with the value "200".
- A text input field for "Перший день" (First day) with the value "2".
- A text input field for "Останній день" (Last day) with the value "100".
- A text input field for "Крок" (Step) with the value "1".
- A text input field for "Поріг" (Threshold) with the value "0.5".

Рис. 6.3. Інтерфейс моделі для введення початкових даних

Для підрахунку необхідно ввести довжину вікна, номер першого та останнього дня проміжку розрахунків, а також поріг для відкидання слабких зв'язків.

Нехай, наприклад, ми маємо базу з 2600 днів. Нам потрібно прорахувати проміжок часу від 2-го до 200-го дня вікнами шириною 200 днів

Довжина вікна повинна бути 200, перший день – 2, останній день – 200.

Для обчислення коефіцієнтів кластеризації необхідно створити так зване дерево найсильніших зв'язків. При цьому ребрами будуть вважатися ті зв'язки, які сильніші (знаходяться ближче один до одного), ніж встановлений поріг. Цей поріг доведеться дібрати, якщо встановлений (0,5) не дасть необхідного результату.

Після введення параметрів та натискання кнопки «Розрахувати» почнеться підрахунок коефіцієнтів. Після підрахунку на екран з'являться вікна з графіками відповідних коефіцієнтів. До речі, після перерахунку коефіцієнтів, ви не зможете відновити старі графіки. Обов'язково збережіть результати обчислень.

Для прикладу розглянемо динаміку зміни коефіцієнтів з часом. На рис. 6.4–6.8 зображено графік коефіцієнту кластеризації, середня довжина шляху, діаметр графу, найвищий степінь вершини та степінь розподілу для індексу MSCI.

Clustering

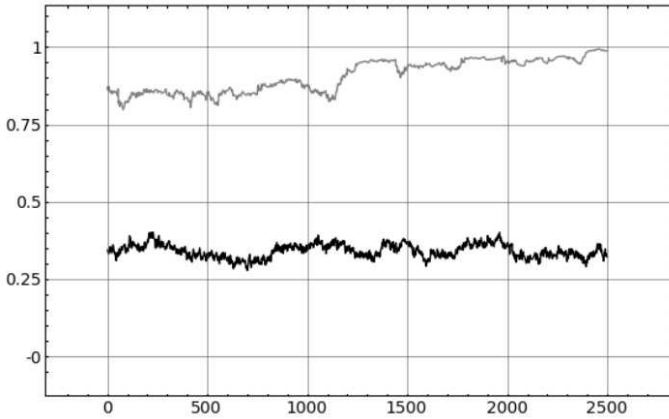


Рис. 6.4. Коефіцієнт кластеризації

Average short path

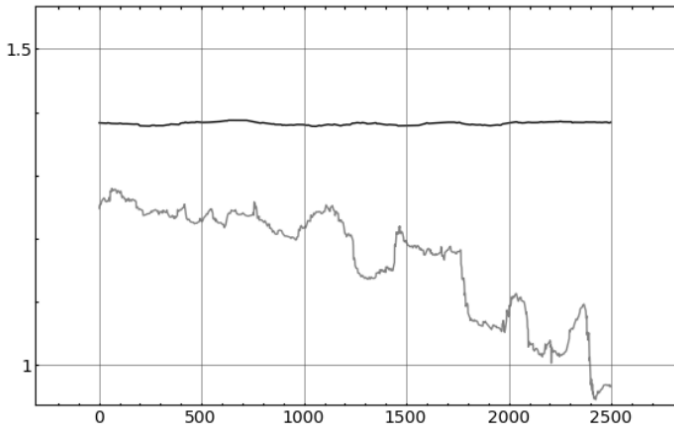


Рис. 6.5. Середня довжина шляху

Diametr

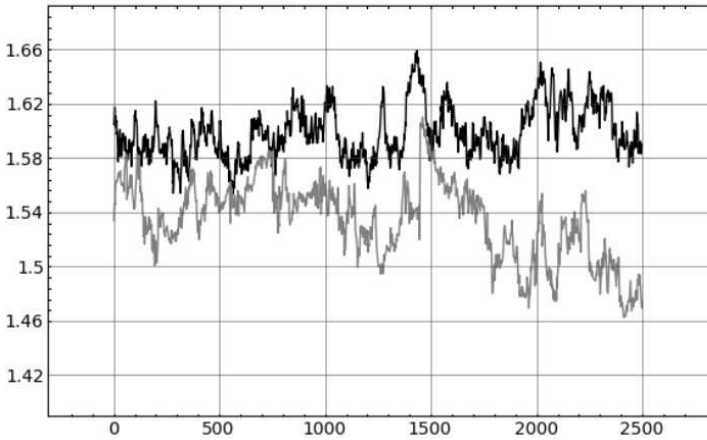


Рис. 6.6. Діаметр графу

Max node degree

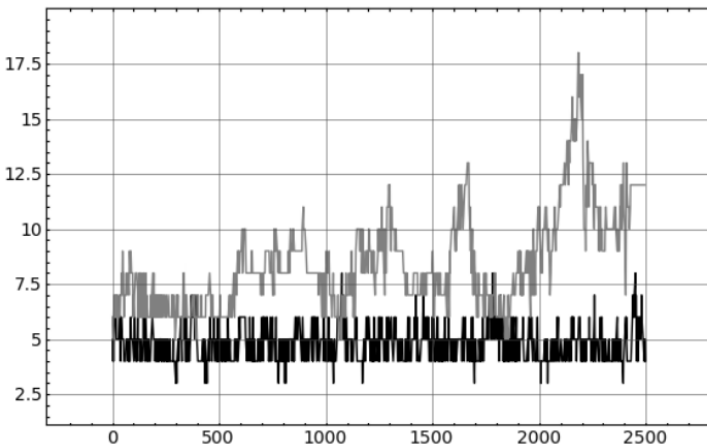


Рис. 6.7. Найвища степінь вершини

На всіх графіках сірим наведені відповідні дані для перемішаної матриці. Видно, що більш складною є система, яка характеризується вихідною матрицею. При перемішуванні складність зникає. Характерна поведінка мір складності кластерних структур дозволить будувати передвісники критичних чи кризових явищ на фінансових ринках.

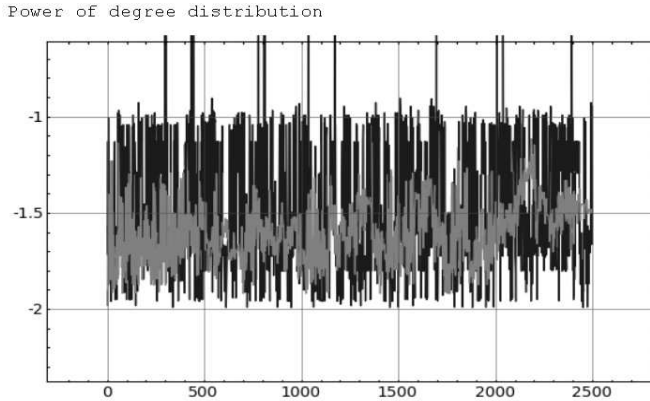


Рис. 6.8. Степінь розподілу

### 6.6. Завдання для самостійної роботи

1. Отримайте модель, базу даних або індекс фонду у викладача та проведіть дослідження:

- коефіцієнта кластеризації;
- середньої довжини шляху;
- діаметра графу;
- найвищої степені вершин.

2. Поясніть і інтерпретуйте результати.

3. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

### 6.7. Контрольні питання

1. Що показують зміни відповідних характеристик графу?

2. Як змінюється коефіцієнт кластеризації під час кризових явищ?

3. Які характеристики графу свідчать про глобалізаційні процеси?

## 7. ВИКОРИСТАННЯ РЕКУРЕНТНОГО АНАЛІЗУ ДЛЯ МОДЕЛЮВАННЯ І ПРОГНОЗУВАННЯ НЕЛІНІЙНИХ ДИНАМІЧНИХ ВЛАСТИВОСТЕЙ СКЛАДНИХ СИСТЕМ

### 7.1. Застосування рекурентного аналізу та рекурентних діаграм до дослідження динаміки та топології складних систем

Дослідження складних систем, як природних, так і штучних, показали, що в їх основі лежать нелінійні процеси, ретельне вивчення яких необхідне для розуміння і моделювання складних систем. У останні десятиліття набір традиційних (лінійних) методик дослідження був істотно розширений нелінійними методами, одержаними з теорії нелінійної динаміки і хаосу; багато досліджень були присвячені оцінці нелінійних характеристик і властивостей процесів, що протікають в природі (скейлінг, фрактальна розмірність). Проте більшість методів нелінійного аналізу вимагає або достатньо довгих, або стаціонарних рядів даних, які досить важко одержати з природи. Більш того, було показано, що дані методи дають задовільні результати для моделей реальних систем, що ідеалізуються. Ці чинники вимагали розробки нових методик нелінійного аналізу даних.

Стан природних або штучних систем, як правило, змінюється в часі. Вивчення цих, часто складних, процесів – важлива задача в багатьох дисциплінах, дозволяє зрозуміти і описати їх суть, наприклад, для прогнозування стану на деякий час в майбутнє. Метою таких досліджень є заходження математичних моделей, які б достатньо відповідали реальним процесам і могли б бути використані для вирішення поставлених задач.

Розглянемо ідею і коротко теорію рекурентного аналізу, наведемо деякі приклади, розглянемо його можливі області застосування при аналізі і прогнозування складних фінансово-економічних систем.

#### *Фазовий простір та його реконструкція*

Стан системи описується її змінними стану

$$x^1(t), x^2(t), \dots, x^d(t), \quad (7.1)$$

де верхній індекс – номер змінної. Набір із  $d$  змінних стану під час  $t$  складає вектор стану  $\vec{x}(t)$  в  $d$ -вимірному фазовому просторі. Даний вектор переміщається в часі в напрямі, визначуваному його вектором

швидкості:

$$\dot{\vec{x}}(t) = \partial_t \vec{x}(t) = \vec{F}(\vec{x}) \quad (7.2)$$

Послідовність векторів  $\vec{x}(t)$  утворює траєкторію у фазовому просторі, причому поле швидкості  $\vec{F}$  дотичне до цієї траєкторії. Еволюція траєкторії описує динаміку системи і її атрактор. Знаючи  $\vec{F}$ , можна одержати інформацію про стан системи в момент  $t$  шляхом інтегрування виразу (7.1). Оскільки форма траєкторії дозволяє судити про характер процесу (періодичні та хаотичні процеси мають характерні фазові портрети), то для визначення стану системи не обов'язково проводити інтегрування, достатньо побудувати графічне відображення траєкторії.

При дослідженні складних систем часто немає інформації про всі змінні стану, або не всі з них можливо виміряти. Як правило, є єдине спостереження, проведене через дискретний часовий інтервал  $\Delta t$ . Таким чином, вимірювання записуються у вигляді ряду  $u_i(t)$ , де  $t = i \cdot \Delta t$ . Інтервал  $\Delta t$  може бути постійним, проте це не завжди можливо і створює проблеми для застосування стандартних методів аналізу даних, що вимагають рівномірної шкали спостережень.

Взаємодії та їх кількість у складних системах такі, що навіть по одній змінній стану можна судити про динаміку всієї системи в цілому. Таким чином, еквівалентна фазова траєкторія, що зберігає структури оригінальної фазової траєкторії, може бути відновлена з одного спостереження або часового ряду за теоремою Такенса методом часових затримок:

$$\hat{\vec{x}}(t) = (u_i, u_{i+\tau}, \dots, u_{i+(m-1)\tau})$$

де  $m$  – розмірність вкладення,  $\tau$  – часова затримка (реальна часова затримка визначається як  $\tau \cdot \Delta t$ ). Топологічні структури відновленої траєкторії зберігаються, якщо  $m \geq 2 \cdot d + 1$ , де  $d$  – розмірність атрактора. На практиці у більшості випадків атрактор може бути відновлений і при  $m \leq 2d$  (рис. 7.1). Затримка, як правило, вибираються апріорно.

Існує декілька підходів до вибору мінімально достатньої розмірності  $m$ , крім аналітичного. Високу ефективність показали методи, засновані на концепції помилкових найближчих точок (false nearest neighbours, FNN). Суть її полягає у тому, що при зменшенні розмірності вкладення відбувається збільшення кількості помилкових точок, що потрапляють в околицю будь-якої точки фазового простору. Звідси витікає простий метод – визначення кількості FNN

як функції від розмірності. Існують й інші методи, засновані на цій концепції – наприклад, визначення відносин відстаней між одними і тими ж сусідніми точками при різних  $m$ . Розмірність атратора також може бути визначена за допомогою крос-кореляційних сум.

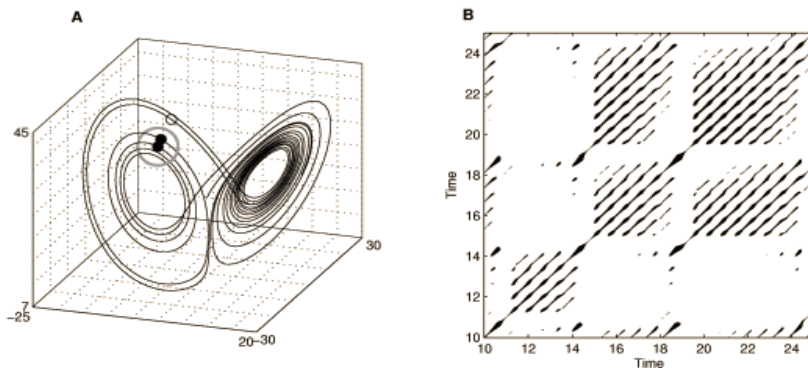


Рис. 7.1. Реконструйований атратор системи Лоренца у тривимірному фазовому просторі і його рекурентна карта

## 7.2. Рекурентний аналіз

Процесам в природі властива яскраво виражена рекурентна поведінка, така, як періодичність або іррегулярна циклічність. Більш того, рекурентність (повторюваність) станів в значенні проходження подальшої траєкторії достатньо близько до попередньої є фундаментальною властивістю дисипативних динамічних систем. Ця властивість була відмічена ще в 80-х роках XIX століття французьким математиком А. Пуанкаре і згодом сформульовано у вигляді теореми рекурентності, опублікованої в 1890 р.

**Якщо система зводить свою динаміку до обмеженої підмножини фазового простору, то система майже напевно, тобто з вірогідністю, практично рівною одиниці, скільки завгодно близько повертається до певного наперед заданого режиму.**

Суть цієї фундаментальної властивості у тому, що, не дивлячись на те, що навіть найменше збурення в складній динамічній системі може привести систему до експоненціального відхилення від її стану, через деякий час система прагне повернутися до стану, деяким чином близького до попереднього, і проходить при цьому подібні етапи еволюції.

Переконалися в цьому можна за допомогою графічного зображе-



ння траєкторії системи у фазовому просторі. Проте можливості такого аналізу сильно обмежені. Як правило, розмірність фазового простору складної динамічної системи більша трьох, що робить практично незручним його розгляд напряду; єдина можливість – проєкції в дво- і тривимірні простори, що часто не дає вірного уявлення про фазовий портрет.

У 1987 р. Ж.-П. Екман [8.4] запропонував спосіб відображення  $m$ -вимірної фазової траєкторії станів системи  $\vec{x}(t)$  завдовжки  $N$  на двовимірну квадратну двійкову матрицю розміром  $N \times N$ , в якій 1 (чорна точка) відповідає повторенню стану при деякому часі  $i$  в деякий інший час  $j$ , а обидві координатні осі є осями часу. Таке подання було назване рекурентною діаграмою, оскільки воно фіксує інформацію про рекурентну поведінку системи.

Математично вищесказане описується як:

$$R_{i,j}^{m,\varepsilon_i} = \Theta(\varepsilon_i - \|\vec{x}_i - \vec{x}_j\|), \vec{x} \in \mathfrak{R}^m, i, j = 1, \dots, N \quad (7.3)$$

де  $N$  – кількість даних станів,  $x_i, \varepsilon_i$  – розмір околиці точки  $\vec{x}$  у момент  $i$ ,  $\|\cdot\|$  – норма і  $\Theta(\cdot)$  – функція Хевісайда.

Непрактично і, як правило, неможливо знайти повну рекурентність у значенні  $\vec{x}_i \equiv \vec{x}_j$  (стан динамічної, а особливо – хаотичної системи не повторюється повністю еквівалентно початковому стану, а підходить до нього скільки завгодно близько). Таким чином, рекурентність визначається як достатня близькість стану  $\vec{x}_j$  до стану  $\vec{x}_i$ . Іншими словами, рекурентними є стани  $\vec{x}_j$ , які потрапляють в  $m$ -вимірну околицю з радіусом  $\varepsilon_i$  і центром в  $\vec{x}_i$ . Ці точки  $\vec{x}_j$  називаються **рекурентними точками** (рис. 7.2).

Оскільки  $R_{i,i} = 1$  ( $i = 1, \dots, N$ ) за визначенням, то рекурентна діаграма завжди містить чорну діагональну лінію – лінію ідентичності під кутом  $\frac{\pi}{4}$  до осей координат. Довільно узята рекурентна точка не несе якої-небудь корисної інформації про стани в часи  $i$  і  $j$ . Тільки вся сукупність рекурентних точок дозволяє відновити властивості системи.

Зовнішній вигляд рекурентної діаграми дозволяє судити про характер процесів, які протікають в системі, наявності і впливі шуму, станів повторення і завмирання (ламінарності), здійсненні в ході еволюції системи різких змін стану (екстремальних подій). На рис. 7.3 наведені деякі з типових процесів і їх рекурентні карти.

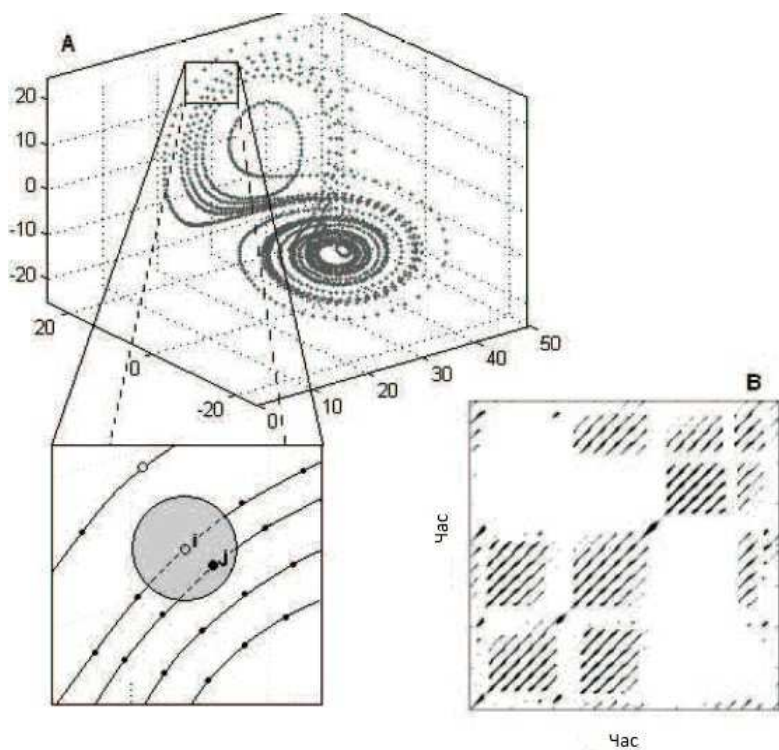
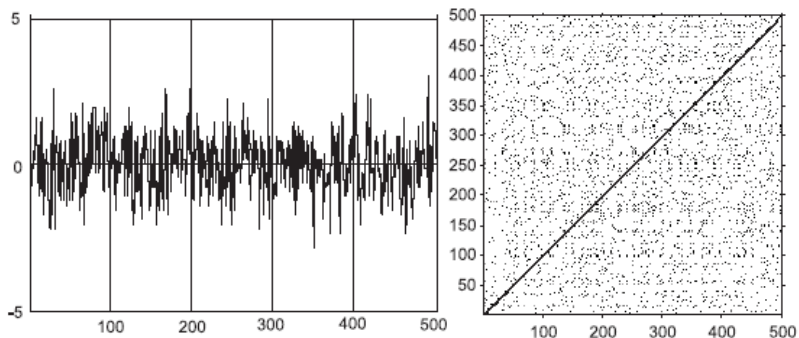
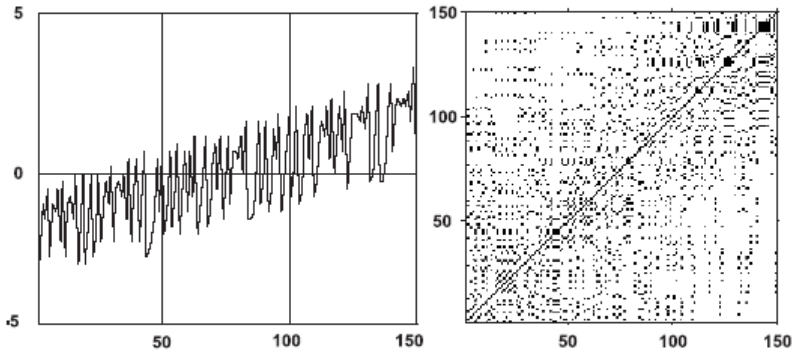


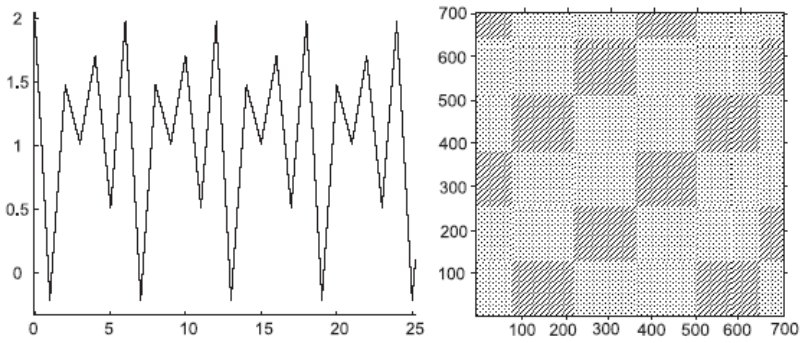
Рис. 7.2. Атрактор Лоренця, вибір рекурентних точок  $i$  та  $j$  в околі  $\varepsilon$ , та рекурентна карта у координатах «час-час»



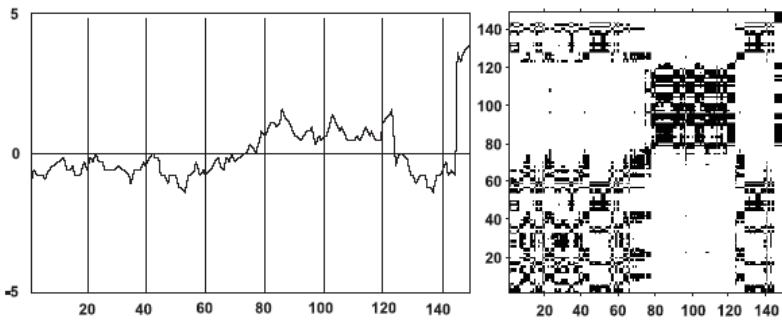
а)



б)



В)



Г)

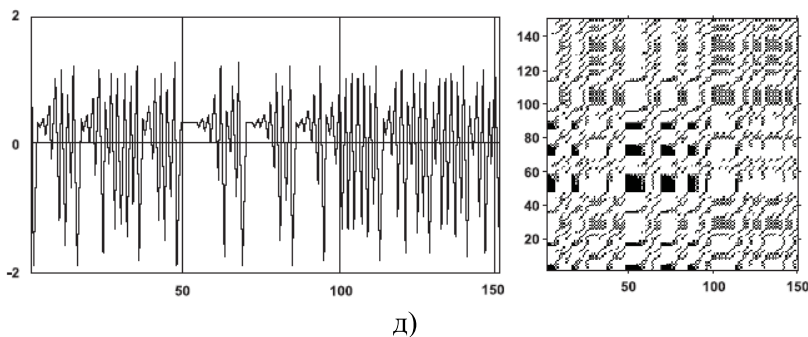


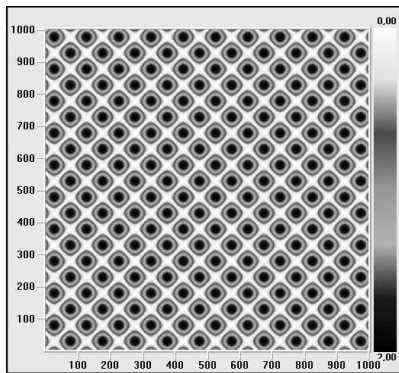
Рис. 7.3. Типові динамічні ряди і їх рекурентні карти:  
 а) однорідна топологія;  
 б) дрейф;  
 в) осцилююча поведінка системи;  
 г) контрастна топологія;  
 д) ламінарність процесу.

Рекурентні діаграми для деяких модельних даних зображені на рис. 7.4.

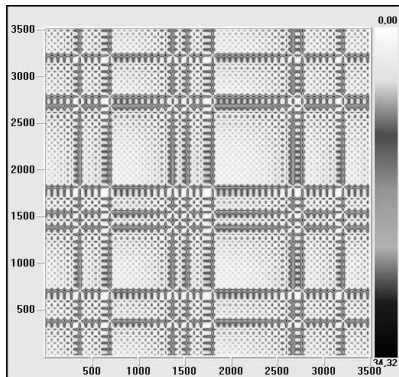
Порогова відстань  $\varepsilon_i$  може вибиратися як безпосередньо для кожної точки, виходячи з отримання в її околі деякої наперед визначеної кількості рекурентних даних точок (при цьому ми одержуємо несиметричну діаграму  $R_{i,j} \neq R_{j,i}$ ), так і залишатися весь час постійним, що дає симетричну діаграму  $R_{i,j} = R_{j,i}$  і є найбільш використовуваним варіантом.

Тип норми також впливає на вигляд діаграми. Як правило, використовуються норми  $L_1$ ,  $L_2$  (евклідова норма),  $L_\infty$  (максимальна норма). Межі цих норм мають різні фігури, що стосовно підходу з постійним  $\varepsilon$  означає, що норма  $L_\infty$  забезпечує знаходження найбільшої,  $L_2$  – середньої, а  $L_1$  – мінімальної кількості сусідніх точок (рис. 7.5). Як правило, при побудові рекурентних діаграм використовується норма  $L_\infty$ , оскільки, по-перше, вона незалежна від розмірності фазового простору, по-друге, найпростіше обчислюється і, по-третє, дозволяє вивчати діаграми теоретично, оскільки аналітичні вирази для  $L_\infty$  розв'язуються простіше, ніж для будь-якої іншої норми. Незалежність від розмірності знаходить застосування при

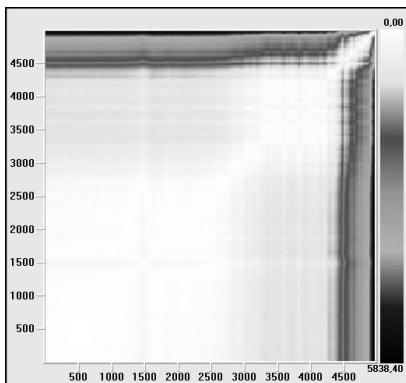
аналізі реконструйованих часових рядів з різними розмірностями вкладення, тобто рекурентні діаграми різних вкладень можуть порівнюватись напряму, тоді як для решти норм потрібне масштабування.



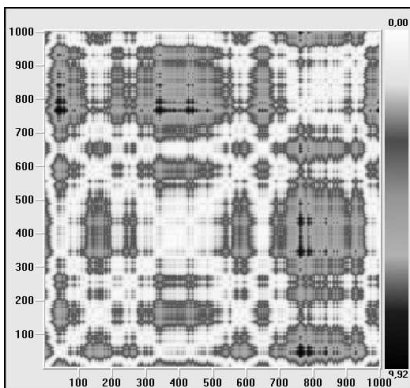
а)



б)



в)



г)

Рис. 7.4. Приклади рекурентних діаграм для:

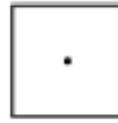
- а) функції  $\sin(x_i)$  ;
- б) атратора Лоренця;
- в) фрагменту фінансового ряду, відомого як індекс Dow Jones;
- г) часового ряду, який відповідає властивостям броунівського руху.



а)



б)



в)

Рис. 7.5. Основні норми з однаковим радіусом навколо точки фазового простору (чорна точка), зображені для двовимірного фазового простору:

а)  $L_1$  ; б)  $L_2$  ; в)  $L_\infty$

### 7.3. Аналіз діаграм

Очевидно, що процеси різної поведінки даватимуть різні рекурентні діаграми. Таким чином, візуальна оцінка діаграм може дати уявлення про еволюцію досліджуваної траєкторії. Виділяють два основних класи структури зображення: топологія, що представляється крупномасштабними структурами, і текстура, що формується дрібномасштабними структурами.

Топологія дає загальне уявлення про характер процесу. Виділяють чотири різних класи (рис. 7.6):

1) однорідні рекурентні діаграми, типові для стаціонарних і автономних систем, в яких час релаксації малий у порівнянні з довжиною ряду;

2) періодичні структури, що повторюються (діагональні лінії, узори у шаховому порядку), відповідають різним осцилюючим системам;

3) дрейф відповідає системам з параметрами, що поволі змінюються, що робить білими лівий верхній і правий нижній кути рекурентної діаграми;

4) різкі зміни в динаміці системи, так само як і екстремальні ситуації, обумовлюють появу білих областей або смуг.

*Рекурентні діаграми спрощують* виявлення екстремальних і рідкісних подій.

Докладний розгляд рекурентних діаграм дозволяє виявити дрібномасштабні структури – текстуру, що складається з простих точок, діагональних, горизонтальних і вертикальних ліній. Комбінації вертикальних і горизонтальних ліній формують прямокутні кла-

стери точок.

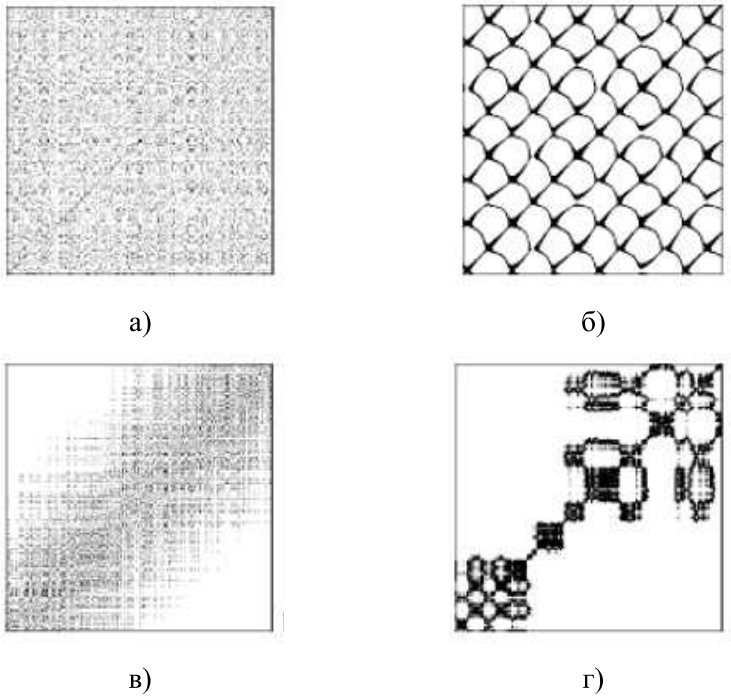


Рис. 7.6. Характерні топології рекурентних діаграм:  
а) однорідна (нормально розподілений шум);  
б) періодична (генератор Ван дер Поля);  
в) дрейф (відображення Ікеди з накладеною послідовністю, що лінійно зростає);  
г) контрастні області або смуги (узагальнений броунівський рух).

Самотні, окремо розташовані рекурентні точки з'являються в тому разі, коли відповідні стани рідкісні, або нестійкі в часі, або викликані сильною флуктуацією. При цьому вони не є ознаками випадковості або шуму.

Діагональні лінії  $R_{i+k, j+k} = 1$  (при  $k = 1 \dots L$ , де  $L$  – довжина діагональної лінії) з'являються у разі, коли сегмент траєкторії у фазовому просторі пролягає паралельно іншому сегменту, тобто траєкторія повторює саму себе, повертаючись в одну і ту ж область фазового простору у різний час. Довжина таких ліній визначається ча-

сом, протягом якого сегменти траєкторії залишаються паралельними; напрям (кут нахилу) ліній характеризує внутрішній час підпроцесів, відповідних даним сегментам траєкторії. Проходження ліній паралельно лінії ідентичності (під кутом  $\frac{\pi}{4}$  до осей координат) свідчить про однаковий напрям сегментів траєкторії, перпендикулярно – про протилежний («відображені» сегменти), що може також бути ознакою реконструкції фазового простору з невідповідною розмірністю вкладення. Нерегулярна поява діагональних ліній є ознакою хаотичного процесу.

*Вертикальні (горизонтальні) лінії*  $R_{i,j+k} = 1$  (при  $k = 1 \dots \nu$ , де  $\nu$  – довжина вертикальної або горизонтальної лінії) виділяють проміжки часу, в котрі стан системи не змінюється або змінюється трохи (система як би «заморожена» на цей час), що є ознакою «ламінарих» станів.

#### 7.4. Хід виконання роботи

Будемо проводити дослідження на робочому аркуші Sage, створеного для побудови рекурентних діаграм.

Виконавши всі комірки, отримаємо початкове вікно, в яке необхідно ввести початкові дані кількість значень та масштаб зображення (рис. 7.7).

Оберіть дію	<input type="text" value="Розрахувати"/>
Кількість значень	<input type="text" value="2000"/>
Масштаб	<input type="text" value="10"/>

Рис. 7.7. Інтерфейс моделі для побудови рекурентної діаграми

Дослідження будемо проводити для часових рядів, які охоплюють період часу, що містить відомі кризи на фінансових ринках. Такими кризами, зокрема, являються кризи 1929 р. («Велика депресія»), 1987 р. («Чорний понеділок») та 2001 р. («криза дот-комів» – криза нової економіки). При цьому часові ряди мають довжину у 2000 днів, а точка кризи знаходиться рівно посередині (точка 1000). Отже, ми можемо прослідкувати динаміку рекурентних характеристик у до кризовий, власне кризовий і після кризовий періоди, виявити спільні і відмінні риси, зробити висновки щодо можливості передбачення кризових явищ.



На рис. 7.8 представлена динаміка індексу Доу Джонса (індекс блакитних фішок на фондовому ринку США), який містить кризу 1929 р. Криза проявляється стрімким падінням індексу у точці 1000.

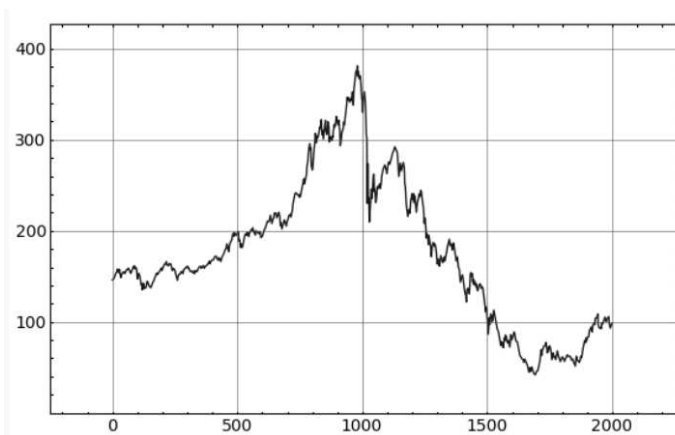


Рис. 7.8. Велика депресія 1929 року на фондовому ринку США

Знайдемо для цього випадку рекурентну карту. Натиснувши «Розрахувати», отримаємо рисунок, подібний зображеному на рис. 7.9.

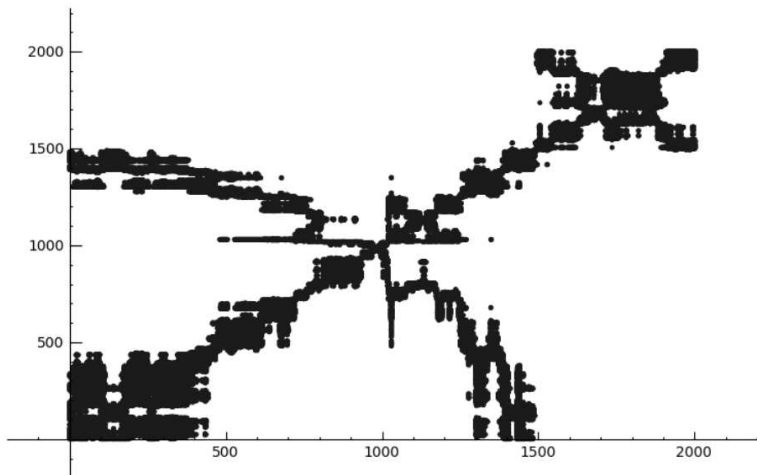


Рис. 7.9. Рекурентна діаграма для часового ряду, що містить кризу 1929 року

Часові ряди, що містять інші кризи та відповідні рекурентні діа-

грами, представлені на рис. 7.10–7.13.

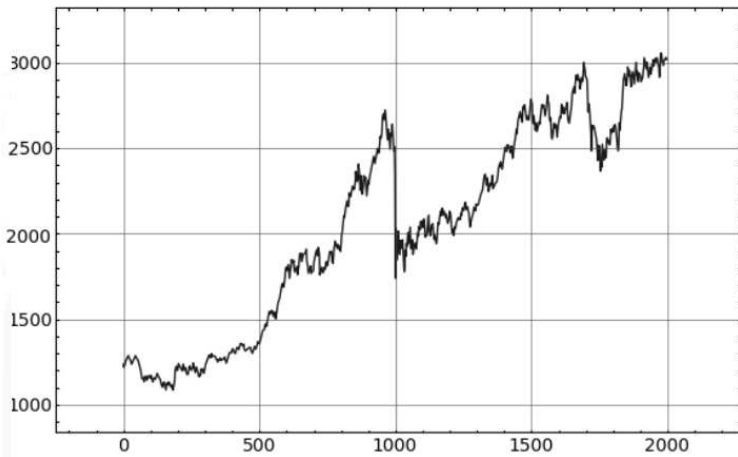


Рис. 7.10. Криза жовтня 1987 року

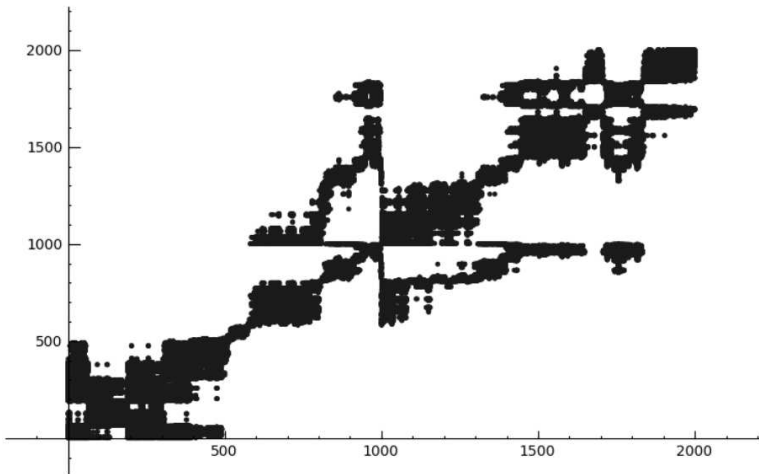


Рис. 7.11. Рекурентна діаграма для кризи 1987 року

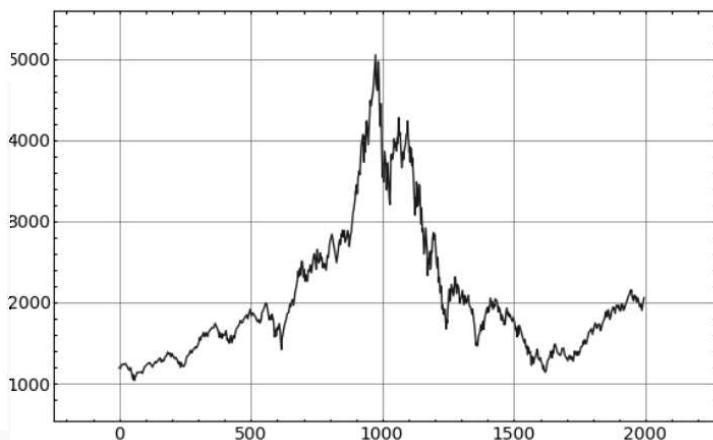


Рис. 7.12. Криза 2001 року за даними високотехнологічного індексу NASDAQ

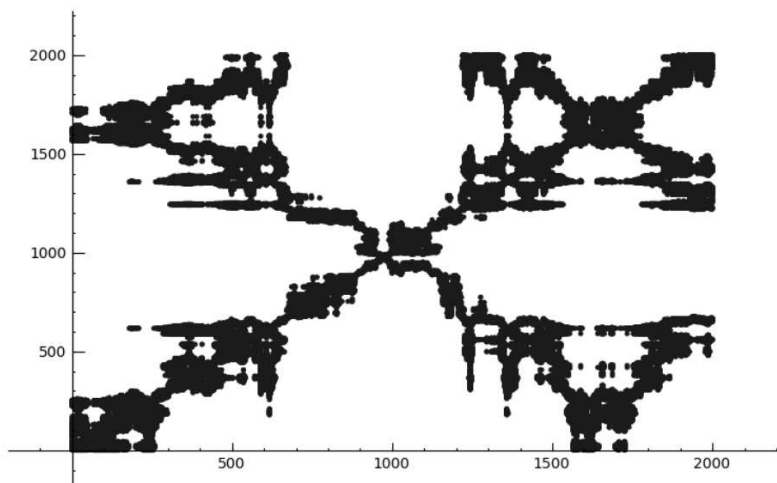


Рис. 7.13. Рекурентна діаграма для кризи нової економіки (весна 2001 року)

Аналіз особливостей рекурентних діаграм дозволяє виявити спільні риси у розподілах як топології, так і структури діаграм. Так, точка кризи є виділеною, а до- та після кризові періоди чітко проявляються на усіх діаграмах. На жаль, візуального аналізу рекурентних діаграм замало для кількісної оцінки стану системи. Тому перейдемо до техніки кількісного аналізу.

### **7.5. Завдання для самостійної роботи**

1. Виберіть за рекомендацією викладача варіанти часових рядів.
2. Проведіть дослідження їх рекурентних властивостей згідно інструкції.
3. Зробіть висновки.
4. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

### **7.6. Контрольні питання**

1. Назвіть сутність фазового простору та його реконструкції.
2. Що таке рекурентні точки.
3. Що допомагають виявити рекурентні діаграми?

## 8. КІЛЬКІСНИЙ АНАЛІЗ РЕКУРЕНТНИХ ДІАГРАМ

### 8.1. Структури в рекурентних графіках

У попередній роботі ми дослідили рекурентні властивості нестационарних часових рядів, виявили і проаналізували дрібномасштабні структури рекурентних діаграм (текстуру, яка складається з простих точок, діагональних, горизонтальних і вертикальних ліній). Комбінації вертикальних і горизонтальних ліній формують прямокутні кластери точок.

Саме ці комбінації використовуються для обчислення мір кількісного аналізу рекурентних діаграм. Дж. А. Збілут і Ч. Л. Веббер [8.4] розробили інструмент обчислення ряду мір на основі підрахунку густини рекурентних точок і побудови частотного розподілу довжин діагональних ліній: RR (коефіцієнт самоподібності), DET (передбачуваність), DIV (зворотна величина максимальної довжини діагональної лінії), ENTR (ентропія) і TREND (тренд). Обчислення цих мір в підматрицях рекурентної діаграми уздовж лінії ідентичності показує поведінку цих мір в часі. Деякі дослідження даних мір показали, що їх застосування може сприяти виявленню точок біфуркації, переходів «хаос—порядок».

Пізніше були запропоновані міри, що використовують густину вертикальних (або горизонтальних) структур: LAM (ламінарність), TT (показник, або час затримки), що дозволяють виявляти переходи «хаос—хаос». Слід зазначити також нечутливість цих мір до вимог стаціонарності і довжини досліджуваних часових рядів.

Уведемо кількісні характеристики введених мір.

Визначимо величини:

$$R_{ij} = \begin{cases} 1, \text{ якщо } (i, j) \text{ рекурентні,} \\ 0, \text{ інакше} \end{cases}$$

$$D_{ij} = \begin{cases} 1, \text{ якщо } (i, j) \text{ та } (i+1, j+1) \text{ або } (i-1, j-1) \text{ рекурентні} \\ 0, \text{ інакше} \end{cases}$$

Тепер коефіцієнти самоподібності і передбачуваності будуть, відповідно, рівні:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{ij}, \quad (8.1)$$

$$DET = \frac{\sum_{i,j=1}^N D_{ij}}{\sum_{i,j=1}^N R_{ij}} \quad (8.2)$$

Якщо  $N_l$  – число діагональних ліній, а  $l_i$  – довжина  $i$ -ої діагональної лінії, то довжина найдовшої діагональної лінії визначається виразом

$$L = \max(l_i; i = 1, \dots, N_l).$$

Ентропія Шеннона для функції  $f(x)$  має вигляд:  

$$H(x) = -\sum_x f(x) \log f(x),$$

а для розподілу:

$$H(p_1, \dots, p_n) = -\sum_{i=1}^n p_i \log(p_i).$$

Для розподілу  $p(l)$  довжини діагональних ліній ентропія Шеннона дорівнює:

$$ENTR = -\sum_{l=1}^{N_l} p(l) \ln p(l), \quad (8.3)$$

Діагональна лінія довжини  $l$  означає, що сегмент траєкторії близький протягом  $l$  кроків часу до іншого сегменту траєкторії в інший час; тому ці лінії пов'язані з дивергенцією сегментів траєкторії. Середня довжина діагональної лінії:

$$L = \frac{\sum_{l=l_{\min}}^N lP(l)}{\sum_{l=l_{\min}}^N P(l)}, \quad (8.4)$$

є середнім часом, що два сегменти траєкторії близькі один одному, і може бути інтерпретований як середній час прогнозу.

Попередні міри відносяться до діагональних (горизонтальних) ліній на рекурентній карті. В той же час існують і вертикальні лінії з відповідними мірами.

Повне число вертикальних ліній довжини  $V$  в RP дається гісто-

грамою:

$$P(v) = \sum_{i,j=1}^N (1 - R_{i,j})(1 - R_{i,j+v}) \prod_{k=0}^{v-1} R_{i,j+k}, \quad (8.5)$$

Аналогічно визначенню детермінізму, відношення рекурентних точок, які формують вертикальні структури до повного набору рекурентних точок може бути обчислено,

$$LAM = \frac{\sum_{v=v_{\min}}^N vP(v)}{\sum_{v=1}^N vP(v)}, \quad (8.6)$$

Ця міра називається **ламiнарнiстю**. Обчислення LAM проводяться для тих  $V$ , які перевищують мінімальну довжину  $v_{\min}$ . Для рекурентних карт часто приймають  $v_{\min} = 2$ . Величина LAM зменшується, якщо RP складається з більш поодиноких рекурентних точок, ніж вертикальні структури.

Середня довжина вертикальних структур дається виразом:

$$TT = \frac{\sum_{v=v_{\min}}^N vP(v)}{\sum_{v=v_{\min}}^N P(v)}, \quad (8.7)$$

і називається **часом затримки** (захвату). Його обчислення вимагає також розгляду мінімальної довжини  $v_{\min}$ , як і у випадку з LAM. TT оцінює середній час, що система перебуватиме в певному стані, або як довго цей стан буде захоплений.

У контрасті до кількісних мір, заснованих на діагональних лініях, введені щойно міри можуть застосовуватись до переходів «хаос-хаос».

Останні два параметри характеризують два різні типи проміжки часу, протягом яких траєкторії перебувають в околі  $\mathcal{E}$ . Їх віконна динаміка дозволяє від слідкувати часову компоненту рекурентних карт.

## 8.2. Хід виконання роботи

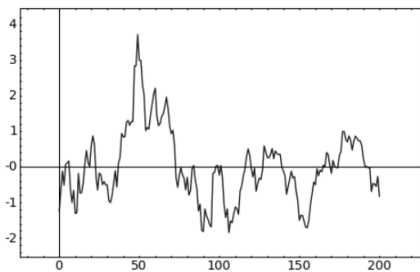
Для проведення кількісного аналізу рекурентних діаграм необхідно завантажити робочий аркуш із підготовленою моделлю (рис. 8.1).

Обчислення пропонується провести для часового ряду,

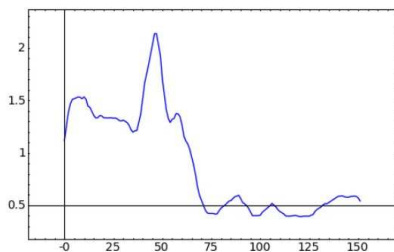
попередньо імпортованого до робочої області з початковим значенням «Розмір вікна=250», «Крок вікна= 2», «Масштаб=5», натиснувши «Рекурентна діаграма» та «Розрахувати» отримаємо рис. 8.2. Перші два графіка зображують відповідно вихідний ряд і його варіацію.

Проведемо тепер дослідження кількісних мір рекурентності для кризових часових рядів. Як і у попередній роботі, розрахуємо типові міри рекурентності і відобразимо разом для порівняння. Результати розрахунків для двох таких мір (міри детермінованості і ентропії) зображені на рис. 8.3–8.4.

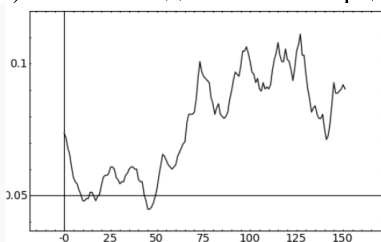
Рис. 8.1. Інтерфейс моделі для проведення дослідження кількісних оцінок рекурентних діаграм



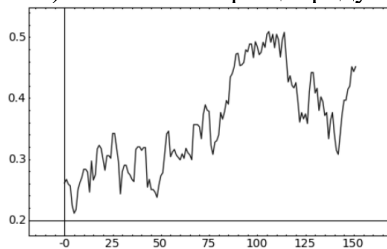
а) Data – вихідний часовий ряд



б) Variance – варіація ряду

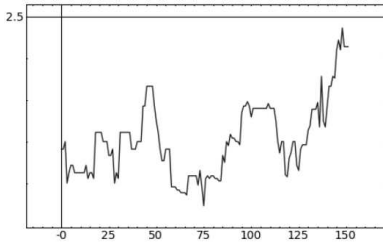


в) RR – коефіцієнт самоподібності

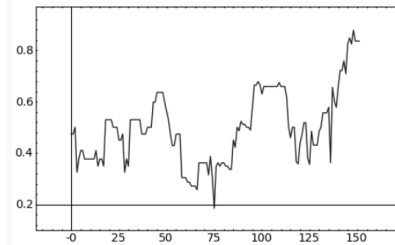


г) DET – коефіцієнт передбачуваності

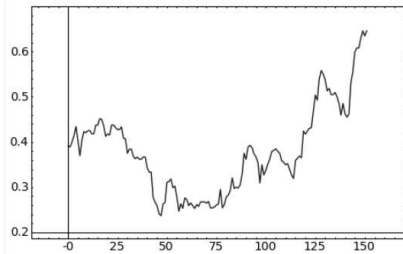




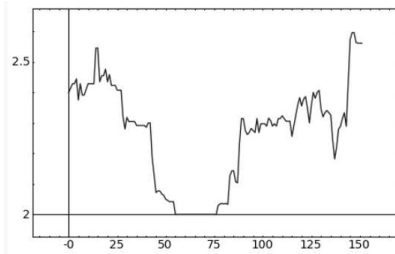
д)  $L$  – середня довжина діагоналей



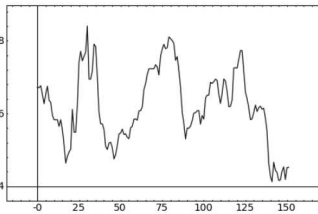
е) ENTR – ентропія



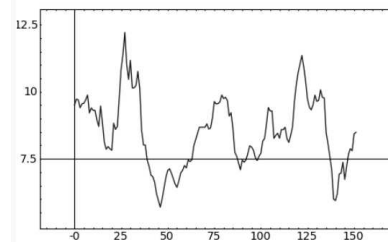
ж) LAM – ламінарність



з) TT – час захвату



к)  $T_1$  – рекурентний час 1-го типу



л)  $T_2$  – рекурентний час 2-го типу

Рис. 8.2. Результати віконних розрахунків мір кількісного аналізу, які дає модель RQA для обраного динамічного ряду

З останніх рисунків видно, що у передкризовий період вказані міри починають суттєво зменшуватись, вказуючи на можливість настання кризи. У після кризовий період ситуація відновлюється.

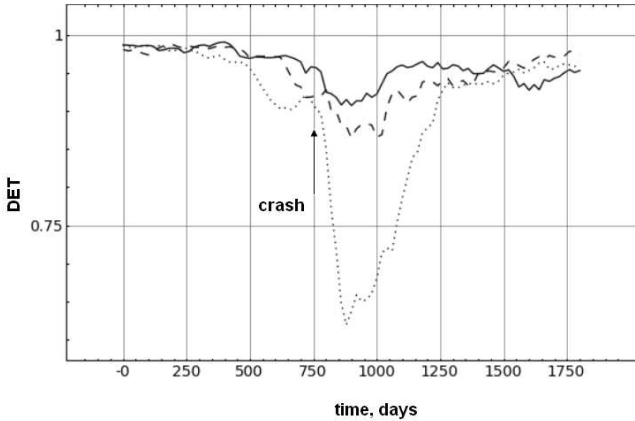


Рис. 8.3. Віконна динаміка міри детермінованості для часових рядів, які містять кризу. Точка початку кризи відмічена стрілкою (crash)

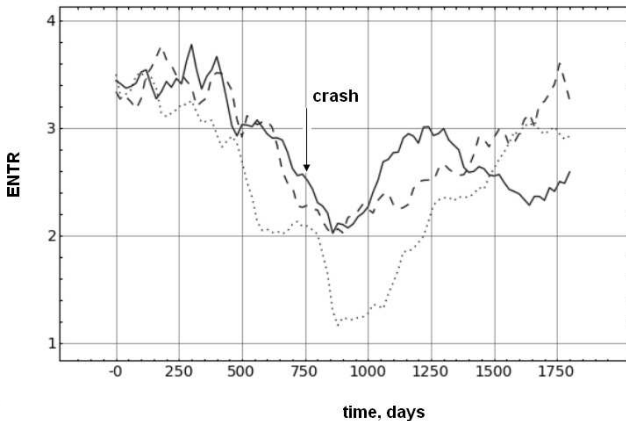


Рис. 8.4. Ентропійна міра рекурентності для кризових часових рядів

Отже, вказані міри (а також і деякі інші) можуть слугувати передвісниками кризових явищ. Зрозуміло, що, досліджуючи післякризову динаміку, можна аналізувати процеси рецесії.

### 8.3. Завдання для самостійної роботи

1. Виберіть за рекомендацією викладача свій варіант часового ряду.
2. Проведіть дослідження динаміки кількісних мір рекурентності

згідно інструкції.

3. Зробіть висновки.

4. Результати-звіт представте у вигляді оформленої індивідуальної роботи.

#### **8.4. Контрольні питання**

1. Назвіть основні кількісних характеристик?

2. За допомогою яких мір визначаються точки біфуркації переходів «хаос-хаос»?

3. Назвіть відмінності діагональних (горизонтальних) мір від вертикальних.

## Додатки

### А. Модуль `finance`

Модуль `finance` було розроблено Вільямом Штейном [8.4] для роботи з фінансовими часовими рядами. Модуль складається з підмодулів `TimeSeries` і `Stock`.

Модуль `TimeSeries` призначено для роботи з дискретними часовими рядами з плаваючою крапкою.

Модуль `Stock` призначено для отримання даних з фондової біржі в реальному часі. В даному класі реалізовані функції для завантаження часових рядів з мережі Інтернет або з деякого файлового ресурсу.

Розглянемо основні методи, які використовуються для отримання характеристик часових рядів:

Методи класу `TimeSeries`

`diffs(...)` – функція повертає новий часовий ряд, отриманий за допомогою знаходження різниці між поточним елементом та попереднім. Таким чином, якщо існує наступний часовий ряд  $X_0, X_1, X_2$ , то за допомогою даної функції отримаємо новий часовий ряд  $X_1 - X_0, X_2 - X_1, \dots$ . Вихідний часовий ряд має на один елемент менше, ніж початковий ряд.

Приклад:

```
sage: v = finance.TimeSeries([5,4,1.3,2,8]); v
[5.0000, 4.0000, 1.3000, 2.0000, 8.0000]
```

```
sage: v.diffs()
[-1.0000, -2.7000, 0.7000, 6.0000]
```

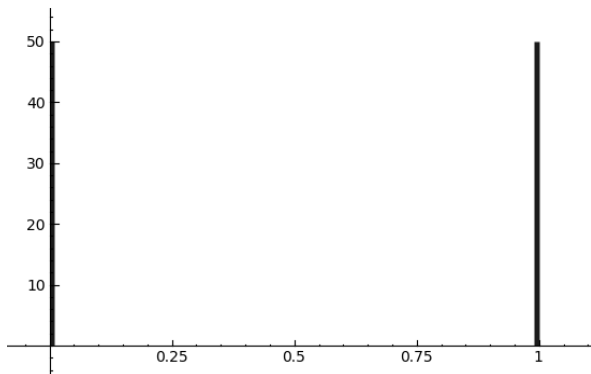
`histogram(...)` – повертає частотну гістограму. Кожний проміжок гістограми представляє кількість попадань значень даної серії в проміжок. Кількість проміжків визначається вхідним параметром `bins` – додатне число (за замовчуванням 50).

```
sage: v = finance.TimeSeries([5,4,1.3,2,8,10,3,-5]); v
[5.0000, 4.0000, 1.3000, 2.0000,
 8.0000, 10.0000, 3.0000, -5.0000]
```

```
sage: v.histogram(3)
[[1, 4, 3], [(-5.0, 0.0), (0.0, 5.0), (5.0, 10.0)]]
```

`plot_histogram()` – повертає зображення гистограми для даного часового ряду із заданою кількістю проміжків. Кількість проміжків визначається вхідним параметром `bins` – додатне число (за замовчуванням – 50).

```
sage: v = finance.TimeSeries([0..1])
sage: v.plot_histogram(bins=100)
```



`sums()` – повертає новий часовий ряд отриманий за допомогою додавання всіх попередніх елементів.

Приклад:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.sums()
[1.0000, 2.0000, 3.0000, 5.0000, 8.0000]
```

`add_scalar()` – повертає новий часовий ряд шляхом додавання скаляру до кожного елемента даної серії.

Вхід функції `s` – число з плаваючою комою.

Приклад:

```
sage: v = finance.TimeSeries([5,4,1.3,2,8,10,3,-5]); v
[5.0000, 4.0000, 1.3000, 2.0000,
8.0000, 10.0000, 3.0000, -5.0000]
```

```
sage: v.add_scalar(0.5)
[5.5000, 4.5000, 1.8000, 2.5000,
8.5000, 10.5000, 3.5000, -4.5000]
```

`simply_moving_average()` – повертає середнє ковзаюче значення часового ряду за останні `k` інтервалів. На `t`-ому кроці виводиться сума своїх попередніх `k-1` кроків та `k`-ий крок поділений

на  $k$ . Таким чином значення  $k$  є середнім до кожної точки.

Вхід функції  $k$  – додатне ціле

Приклад:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v  
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.simple_moving_average(0)  
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.simple_moving_average(1)  
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.simple_moving_average(2)  
[1.0000, 1.0000, 1.0000, 1.5000, 2.5000]
```

```
sage: v.simple_moving_average(3)  
[1.0000, 1.0000, 1.0000, 1.3333, 2.0000]
```

`standart_deviation()` – повертає стандартне відхилення вхідної послідовності.

Вхід функції – логічний параметр `bias` (за замовчуванням: `False`), якщо `False` відбувається ділення на `length()` – 1, замість `length()`, щоб дати більш об'єктивну оцінку дисперсії.

Вихід – дійне число.

Приклад:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v  
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.standard_deviation()  
0.8944271909...
```

```
sage: v.standard_deviation(bias=True)  
0.8000000000...
```

`mean()` – повертає середнє значення із елементів послідовності.

Приклад:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v  
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.mean()  
1.6000000000000001
```

`exponential_moving_average()` – повертає експоненційне ковзне середнє часового ряду. На  $t$ -му кроці виводиться

сума попередніх  $k-1$  кроків та на  $k$ -му кроці відбувається ділення на  $k$ .

Вхід функції –  $\alpha$  – дійсне значення від 0 до 1.

**Приклад:**

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v  
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.exponential_moving_average(0)  
[0.0000, 1.0000, 1.0000, 1.0000, 1.0000]
```

```
sage: v.exponential_moving_average(1)  
[0.0000, 1.0000, 1.0000, 1.0000, 2.0000]
```

```
sage: v.exponential_moving_average(0.5)  
[0.0000, 1.0000, 1.0000, 1.0000, 1.5000]
```

```
randomize(distribution, loc='uniform',  
scale='0', **kws='1') – генерує випадковий часовий ряд
```

**Параметри функції:**

`distribution` –

`uniform` – рівномірно в інтервалі від `loc` до `loc+scale`;

`normal` – середнє значення `loc` із стандартним відхилення

`scale`

`loc` – дійсне (за замовчуванням 0);

`scale` – дійсне (за замовчуванням 1).

**Приклад:**

Згенеруємо 5 рівномірно випадкових чисел в інтервалі [0;1]

```
sage: set_random_seed(0)  
sage: finance.TimeSeries(5).randomize()  
[0.8685, 0.2816, 0.0229, 0.1456, 0.7314]
```

Згенеруємо 5 рівномірно випадкових чисел в інтервалі [5;7]

```
Sage: set_random_seed(0)  
sage: finance.TimeSeries(10).\randomize('uniform', 5, 2)  
[6.7369, 5.5632, 5.0459, 5.2911, 6.4628,  
5.2412, 5.2010, 5.2761, 5.5813, 5.5439]
```

Згенеруємо 5 нормально випадкових чисел із середнім 5 та різницею 2:

```
sage: set_random_seed(0)  
sage: finance.TimeSeries(10).\
```

```
randomize('normal', 5, 2)
[6.3534, 4.1978, 5.7153, 3.8273, 3.1269,
2.9598, 3.7484, 6.7472, 3.8986, 4.6271]
```

`variance ()` – повертає дисперсію елементів поточного часового ряду, що представляють собою середній квадрат відхилення від середнього значення.

Параметр функції – `bias` – логічний (за замовчуванням `False`), якщо `False` ділимо на `length () - 1` замість `length ()` для отримання менш упередженої дисперсії

Приклад:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.variance()
0.800000000000000004
```

```
sage: v.variance(bias=True)
0.640000000000000001
```

`hurst_exponent ()` – обчислення показника Херста часового ряду.

Приклад:

```
sage: set_random_seed(0)
sage: bm = finance.TimeSeries(10^5).\
randomize('normal').sums(); bm
[0.6767, 0.2756, 0.6332, 0.0469, -0.8897 ...
152.2437, 151.5327, 152.7629, 152.9169,
152.9084]
```

```
sage: bm.hurst_exponent()
0.527450972...
```

`plot ()` – функція повертає зображення часового ряду у вигляді лінії чи точок  $(i, T(i))$ , де  $i$  – невід’ємний діапазон цілих чисел від 0 до довжини часового ряду.

Параметри функції

`plot_points` – (за замовчуванням 1000) 0 або додатне ціле; будується лише зображення заданої кількості рівновіддалених точок в часовому ряді, якщо 0, то зображуються всі точки

`points` – логічний (по замовчуванню `False`), якщо `True`, повертаються лише точки часового ряду.

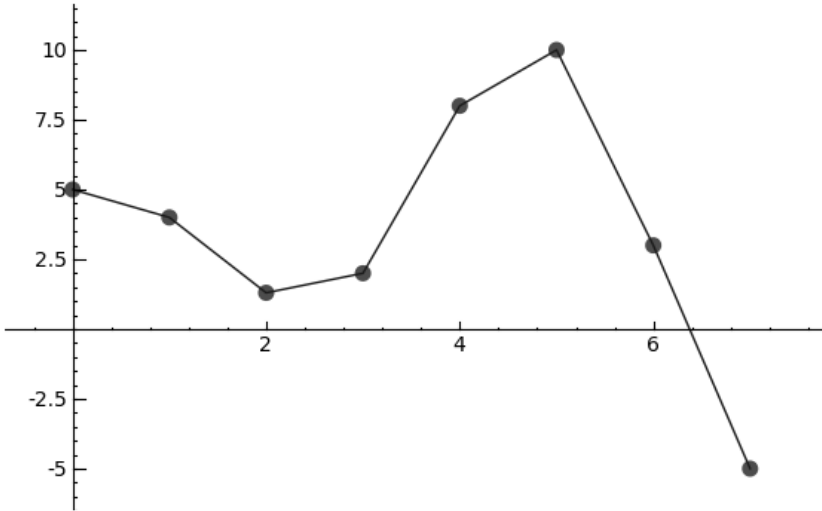
`**kws` – додаткові параметри зображення ліній або точок.



Приклад:

```
sage: v = finance.TimeSeries([5,4,1.3,2,\
      8,10,3,-5]); v
[5.0000, 4.0000, 1.3000, 2.0000,
 8.0000, 10.0000, 3.0000, -5.0000]

sage: v.plot() + v.plot(points=True,\
      rgbcolor='red',pointsize=50)
```



`exp()` – повертає новий часовий ряд отриманий за допомогою застосування експоненціального відображення для всіх членів ряду. В результаті отримуємо новий часовий ряд.

Приклади:

```
sage: v = finance.TimeSeries([1..5]); v
[1.0000, 2.0000, 3.0000, 4.0000, 5.0000]

sage: v.exp()
[2.7183, 7.3891, 20.0855, 54.5982, 148.4132]

sage: v.exp().log()
[1.0000, 2.0000, 3.0000, 4.0000, 5.0000]
```

`correlation()` – повертає кореляцію поданого часового ряду, що є коваріацією, поділену на добуток їх стандартних відхилень. Параметри функції:

other – часовий ряд

Приклади:

```
sage: v = finance.TimeSeries([1,-2,3]);
      w = finance.TimeSeries([4,5,-10])

sage: v.correlation(w)
-0.558041609...

sage: v.covariance(w)/(v.standard_deviation()
      * w.standard_deviation())
-0.558041609...
```

`autocorrelation()` – повертає кореляцію часового ряду самого з собою (між попередніми та наступними значеннями). Автокореляцію іноді називають послідовною кореляцією, що означає кореляцію між членами ряду чисел, розташованих у певному порядку. Також синонімами цього терміну є лагова кореляція та персистентність.

Параметр функції –  $k$  – невід’ємне ціле число.

Приклади:

```
sage: v = finance.TimeSeries([13,8,15,4,\
      4,12,11,7,14,12])

sage: v.autocorrelation()
-0.1875

sage: v.autocorrelation(1)
-0.1875

sage: v.autocorrelation(0)
1.0

sage: v.autocorrelation(2)
-0.20138888888888887

sage: v.autocorrelation(3)
0.18055555555555555

sage: finance.TimeSeries([1..1000]).\
      autocorrelation()
0.997
```

`central_moment()` – повертає центральний момент  $k$ -го порядку випадкової величини  $\xi$ , що є математичним сподіванням

величини.

Параметр функції –  $k$  – додатне число.

Приклад:

```
sage: v = finance.TimeSeries([1,2,3])
```

```
sage: v.central_moment(2)
0.66666666666666663
```

`moment()` – повертає початковий момент  $k$ -го порядку випадкової величини  $\xi$ , який є математичним сподіванням величини.

Параметр функції –  $k$  – додатне число.

Приклади:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.moment(1)
1.60000000000000001
```

```
sage: v.moment(2)
3.20000000000000002
```

`max()` – повертає максимальне значення часового ряду.

Приклади:

```
sage: v = finance.TimeSeries([1,-4,3,-2.5,-4,3])
```

```
sage: v.max()
3.0
```

```
sage: v.max(index=True)
(3.0, 2)
```

`min()` – повертає мінімальне значення часового ряду.

Приклади:

```
sage: v = finance.TimeSeries([1,-4,3,-2.5,-4])
```

```
sage: v.min()
-4.0
```

```
sage: v.min(index=True)
(-4.0, 1)
```

`mean()` – повертає середнє значення часового ряду.

Приклади:

```
sage: v = finance.TimeSeries([1,1,1,2,3]); v
[1.0000, 1.0000, 1.0000, 2.0000, 3.0000]
```

```
sage: v.mean()
1.6000000000000001
```

`range_statistic()` – повертає інтервал, який представляє собою різницю між максимальним та мінімальним значенням цього ряду, поділене на стандартне відхилення різниці між цими елементами.

Приклади:

```
sage: set_random_seed(0);
      finance.TimeSeries(10^6).\
      randomize('normal').sums().\
      range_statistic()
1897.8392602
```

```
sage: set_random_seed(0);
      finance.TimeSeries(10^6).\
      randomize('normal',0,100).sums().\
      range_statistic()
1897.8392602...
```

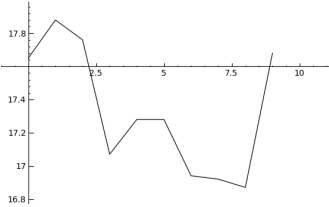
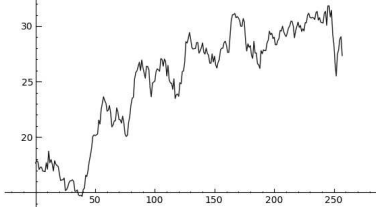
Отримати повний список допустимих методів для опрацювання часових рядів можна за вказівкою `finance.TimeSeries.<Tab>`.

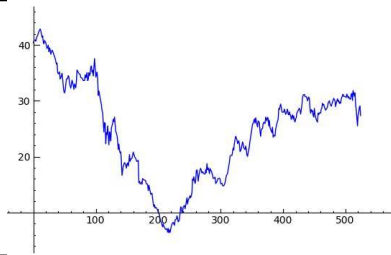
## Модуль Stock

Модуль призначений для отримання інформації з фондової біржі. Даний модуль дозволяє отримувати дані з фондової біржі в реальному часі. В даному класі реалізовані функції для завантаження часових рядів з мережі Інтернет, або ж з деякого файлового ресурсу.

`close()` – значення ціни на момент закриття торгів для індексу або акції.

Приклади:

	
<pre>sage: # останні 10 днів cc=finance.Stock('dow').\ close()[ :10].plot().show()</pre>	<pre>sage: #за минулий рік cc=finance.Stock('dow').\ close().plot().show()</pre>



```
sage: # з 1 Травня 2008 року по теперешній час
cc=finance.Stock('dow').close('May+1,+2008').plot().show()
```

`open()` – значення ціни на момент відкриття торгів для індексу або акції.

Приклад:

```
sage: cc=finance.Stock('dow').open()
```

`google()` – дані торгів для акцій або фонду протягом дня, зазначених у `Stock`, отриманих від Google:

Приклад:

```
sage: finance.Stock('vmw').google('May+10,+2010')[ :10]
[
10-May-10 59.66 61.19 59.66 60.64      2295165,
11-May-10 59.83 62.62 59.50 60.87      1968566,
12-May-10 61.40 62.92 61.19 62.84      1745673,
13-May-10 62.61 64.59 61.78 62.19      3097101,
14-May-10 61.16 61.38 58.93 60.20      2434423
]
```

`yahoo()` – поточні дані фонду або акцій, отримані з Yahoo Finance.

Приклад:

```
sage: finance.Stock('ibm').yahoo()
{'stock_exchange': 'NYSE', 'market_cap': '168.2B',
'200day_moving_avg': '127.595', '52_week_high': '134.25',
'price_earnings_growth_ratio': '1.07', 'price_sales_ratio':
'1.74',
'price': '131.19', 'earnings_per_share': '10.29',
'50day_moving_avg':
'129.014', 'avg_daily_volume': '6818940', 'volume':
'9920567',
'52_week_low': '99.50', 'short_ratio': '1.40',
'price_earnings_ratio':
'12.78', 'dividend_yield': '1.75', 'dividend_per_share':
'2.30',
```

```
'price_book_ratio': '7.65', 'ebitda': '24.033B', 'change':  
'-0.29',  
'book_value': '17.182'}
```

`market_value()` – поточна ціна акції або поточне значення фонду.

Приклад:

```
sage: finance.Stock('aapl').market_value()  
253.81999999999999
```

```
sage: finance.Stock('dow').market_value()  
27.34
```

`historical()` – повертає незмінну послідовність цінових даних для даного фондового ринку, отриманих від Google.

Параметр функції `startdate` – рядок (за замовчуванням `'Jan+1,+1990'`)

Приклад:

Отримаємо історію торгів за останні п'ять днів компанії VMware

```
sage: finance.Stock('vmw').historical()[:5]  
[  
15-Aug-07 52.11 59.87 51.50 57.71 10678500,  
16-Aug-07 60.99 61.49 52.71 56.99 6919500,  
17-Aug-07 59.00 59.00 54.45 55.55 3086100,  
20-Aug-07 56.05 57.50 55.61 57.33 2140900,  
21-Aug-07 57.25 66.59 56.50 65.99 7369700  
]
```

## Б. Модуль Graph

Створення та опрацювання графів у Sage можна здійснювати за допомогою: а) убудованих засобів (убудованих модуля `Sage.graphs.graph` та класу `graphs`); б) зовнішніх пакетів (бібліотеки `Networkx`; команд `Maxima`, `GAP`, `Magma`, `Maple`, `Mathematica` тощо).

У ММС Sage універсальним можна вважати спосіб створення неорієнтованого графу за допомогою функції `Graph([arguments])`, що являє собою конструктор однойменного класу. Звернення до функції може бути виконане в одному із трьох форматів: *скороченому*, *основному* та *розширеному*.

При *скороченому* зверненні (без зазначення аргументів) – `G1=Graph()` – буде створено порожній граф `G1`. Звернення до функції в *основному* форматі – `Graph(data)` – є необхідним для ви-

значення графу певної структури, наприклад: за вказівкою  $G2=Graph(\{1:[]\})$  буде створено граф  $G2$  з однією нумерованою вершиною,  $G3=Graph(\{0:[],1:[],2:[]\})$  – граф  $G3$ , що має три нумеровані вершини,  $G4=Graph('A':[],'B':[],'C':[]\})$  – граф  $G4$ , що має три іменовані вершини,  $G5=Graph(\{0:[3],1:[2,0],3:[1]\})$  – граф  $G5$ , що має чотири нумеровані вершини (0, 1, 2, 3) та чотири ребра – (0,3), (1,2), (1,0) і (3,1) (рис. Б. 1).

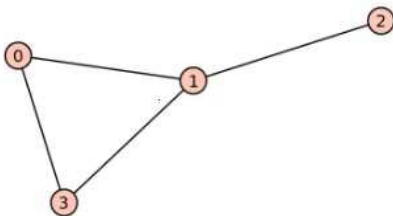


Рис. Б.1. Граф, створений за вказівкою  $G5=Graph(\{0:[3],1:[2,0],3:[1]\})$

Окрім зазначеного у прикладах формату подання структури графу – як словника бібліотеки `NetworkX` – допускається задання параметру `data` у вигляді словника словників, словника списків, матриці у форматі `NumPy` та `ndarray`, матриці інцидентності або матриці суміжності (табл. Б.1).

Таблиця Б.1.

**Формати задання параметру `data`**

Формат	Приклад застосування
Словник словників	<code>g=Graph(\{0:\{1:'x',2:'z',3:'a'\},2:\{5:'out'\}\})</code>
Словник списків	<code>g=Graph(\{0:[1,2,3], 2:[4]\})</code>
Матриця NumPy	<code>import numpy g=Graph(numpy.array([[0,1,1,1,0], [1,0,0,0,0],[1,0,0,0,1],[1,0,0,0,0], [0,0,1,0,0]]))</code>
Матриця інцидентності	<code>g=Graph(matrix([(0,1,1,1,0),(1,0,0,0,0), (1,0,0,0,1),(1,0,0,0,0),(0,0,1,0,0)]))</code>
Матриця суміжності	<code>g=Graph(Matrix([(-1,-1,-1,0),(0,0,1,0), (0,1,0,-1),(1,0,0,0),(0,0,0,1)]))</code>

Під *розширеним* форматом звернення до функції `Graph()` розуміється таке звернення, при якому крім основного аргументу

(data) зазначаються додаткові аргументи за правилом: `назва_аргументу=значення` (табл. Б.2). Порядок зазначення аргументів не є строгим.

**Таблиця Б.2**

**Додаткові аргументи функції `Graph()`**

Назва аргументу	Призначення аргументу
<code>name</code>	ім'я графу
<code>loops</code>	ознака наявності петель у графі
<code>multiedges</code>	ознака зображення всіх ребер між парою вершин
<code>weighted</code>	ознака зваженості графу
<code>vertex_labels</code>	ознака наявності/відсутності назви вершин
<code>sparse</code>	ознака розрідженості графу

Окрім універсального способу створення графів, у Sage існує набір функцій класу `graphs` для конструювання окремих груп графів, наприклад:

– за виглядом: `Grid2dGraph()`, `FlowerSnark()`, `TetrahedralGraph()` та ін.;

– за способом формування: `RandomGNP()`, `RandomRegular()`, `RandomShell()`;

– «іменні» графи: `PetersonGraph()`, `HoffmanSingletonGraph()`, `MobiusKantorGraph()` та ін.;

– сімейства графів: `CirculantGraph()`, `CubeGraph()`, `BalancedTree()` та ін.

Побудова та зображення графів на площині та у просторі виконується за допомогою функцій `plot()` та `plot3d()`, `show()` та `show3d()` відповідно (рис. Б.2).

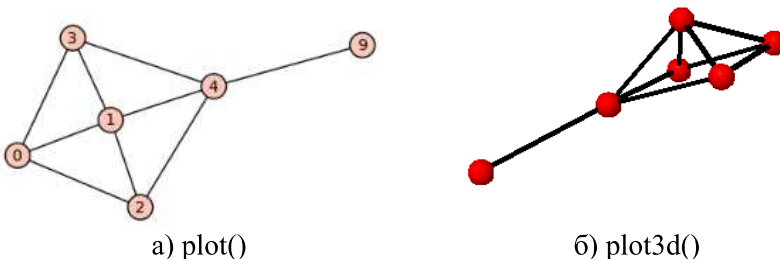


Рис. Б.2. Побудова графу на площині (а) та у просторі (б)



У таблицях Б.3 та Б.4 подано перелік та призначення основних параметрів функцій `plot()` та `show()`.

**Таблиця Б.3**

**Параметри функції `plot()`**

Назва параметру	Призначення параметру
<code>layout</code>	спосіб розташування вершин графу ( <code>circular</code> – зображення вершин графу у вигляді кола, <code>tree</code> – зображення графу у вигляді дерева; корінь дерева задається параметром <code>tree_root</code> , в протилежному випадку корінь буде обрано у випадковий спосіб)
<code>vertex_labels</code>	ознака виведення назв вершин
<code>vertex_size</code>	розмір вершин (за замовчуванням 200)
<code>vertex_shape</code>	формат зображення вершин ('^' – у вигляді трикутника, 'o' – у вигляді кола, 's' – у вигляді квадрату, 'p' – у вигляді п'ятикутника тощо; за замовчуванням 'o')
<code>vertex_color</code>	колір вершин
<code>graph_border</code>	ознака наявності/відсутності рамки навколо графу
<code>edge_labels</code>	ознака виведення назв ребер
<code>edge_colors</code>	колір ребер
<code>edge_style</code>	стиль ребер ( <code>solid</code> – суцільна лінія, <code>dashed</code> – штрих-пунктирна лінія, <code>dotted</code> – пунктирна лінія)
<code>partition</code>	ознака зображення ребер різними кольорами
<code>color_by_label</code>	колір назв
<code>tree_orientation</code>	розміщення кореня дерева (догори, донизу)

**Таблиця Б.4**

**Параметри функції `show()`**

Назва параметру	Призначення параметру
<code>frame</code>	ознака наявності/відсутності рамки навколо зображення (за замовчуванням відсутня)
<code>fontsize</code>	розмір шрифту, додатне ціле число
<code>figsize</code>	розмір вершини, додатне ціле число
<code>xmin</code>	мінімальна координата по вісі x
<code>xmax</code>	максимальна координата по вісі x

ymin	мінімальна координата по вісі y
ymax	максимальна координата по вісі y
gridlines	сітка з кроком figsize
aspect_ratio	коефіцієнт пропорційності зображення

Виконання різноманітних операцій над елементами графу передбачає звернення до відповідних методів класу Graph. Перелік та призначення основних методів подано в таблиці Б.5. Отримати повний список допустимих методів для опрацювання графів можна за вказівкою Graph.<Tab>

**Таблиця Б.5**

**Основні методи опрацювання графів**

<b>Назва методу</b>	<b>Призначення методу</b>
edges(labels, sort)	визначення списку ребер, кожне з яких подається трійкою чисел (u,v,l), де u та v – вершини графу, l – назва ребра
vertices()	визначення списку вершин
add_edge(u, v, label)	додавання ребра label, що з'єднує вершини u та v
add_vertex(name)	додавання ізольованої вершини з іменем name
add_vertices(vertices)	додавання вершин, що задаються списком vertices
delete_vertex(vertex, in_order)	видалення вершини з іменем vertex
set_edge_label(u, v, l)	надання ребру (u,v) імені – l
add_edges(edges)	додавання ребер за списком edges
num_edges()	визначення кількості ребер графу
num_verts()	визначення кількості вершин графу
degree(vertices, labels)	визначення степеня вершини vertices
name(new)	перейменування графу
shortest_path(u, v, by_weight, bidirectional)	визначення найкоротшого шляху від вершини u до вершини v у вигляді списку вершин
all_paths(start, end)	визначення списку всіх шляхів

Назва методу	Призначення методу
	між парою вершин (start,end)
<code>distance(u,v)</code>	визначення відстані (довжини найкоротшого шляху) між парою вершин (u,v)
<code>min_spanning_tree()</code>	визначення мінімального остівного дерева
<code>clustering_coeff(nbunch, with_labels, weights)</code>	визначення списку коефіцієнтів кластеризації для кожної вершини зі списку nbunch у вигляді впорядкованого списку
<code>clustering_average()</code>	визначення середнього коефіцієнту кластеризації графу
<code>cluster_triangles(nbunch, with_labels)</code>	визначення кількості трикутників для вершин упорядкованого списку nbunch
<code>cluster_transitivity()</code>	визначення коефіцієнту транзитивності

Окрім командного способу звернення до функцій перетворення графу, у Sage передбачена можливість подання графу у редакторі з напівавтоматичним режимом управління. Відкриття редактору здійснюється за вказівкою `graphs_editor(graph)`, наприклад (рис. Б.3):

```
sage : g=graphs.CompleteGraph(3)
graph_editor(g)
```

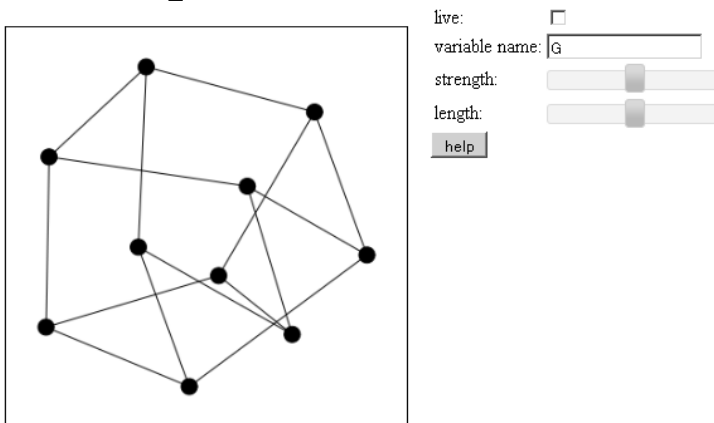


Рис. Б.3. Інтерфейс редактору графів MMC Sage

У редакторі над графом можуть бути виконані такі операції: вибір, переміщення та видалення будь-якої вершини; створення нової вершини; створення або видалення ребер тощо.

## В. Модуль dynamics

Розроблений Н. А. Хараджян модуль `dynamics` призначено для побудови траєкторій руху динамічних систем фрактальних об'єктів включає наступні функції: `evolution`, `evolution_2d`, `chaosgame`, `staircase`, `ifs`, `orbits` та `rk`.

Функція `evolution(fun, initial, n)` розраховує множину  $n+1$  точок, які зобразимо на двовимірному графіку, де горизонтальні координати точок є числа 0, 1, 2, 3 і т.д., а вертикальні координати – відповідні значення послідовності, визначені рекурентним співвідношенням  $fun(y_{n+1}=F(y_n))$  з початковим значенням `initial`.

Для зображення послідовності точок  $3, \cos(3), \cos(\cos(3)), \dots$  виконаємо наступні команди (рис. В.1):

```
sage: var('y')
      evolution(cos(y), 2, 20);
      show(point(a, pointsize=5))
      cos(y)
```

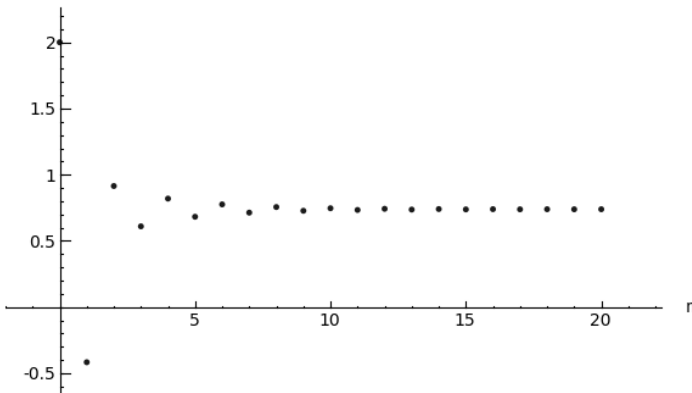


Рис. В.1. Графічне подання послідовності точок  $3, \cos(3), \cos(\cos(3)), \dots$

Функція `evolution_2d(fun, state, initial, n)` розраховує розташування на площині  $(n+1)$ -ої точки послідовності, що визначається дискретною динамічною системою з рекурентними співвідношеннями  $fun(u_{n+1}=F(u_n, v_n), v_{n+1}=G(u_n, v_n))$  з змінними положення

state та початковими значеннями initial ( $u_0, v_0$ ).

Приклад: дискретні двовимірні системи, що мають хаотичні розв'язки, називаються картами Хенона та визначають за допомогою рекурентного співвідношення з параметрами  $a$  і  $b$ :

$$y_{n+2} = 1 + ay_n + by_{n+1}^2$$

Для подання карт Хенона у вигляді системи рівнянь першого порядку введемо підстановку  $y_{n+1} = x_n$ :

$$\begin{cases} x_{n+1} = 1 + ay_n + bx_n^2 \\ y_{n+1} = x_n \end{cases}$$

Отримуємо систему функцій  $\{x_{n+1}, y_{n+1}\}$  відносно змінних  $x, y$ .

При різних значеннях  $a$  і  $b$  можна одержати різні траєкторії руху динамічної системи:

```
sage: a=evolution_2d([1+0.3*y-1.4*x^2,x], [x,y],
[0.2,0.2], 100000) #задання моделі та обчислювальний експеримент
show(point(a,pointsize=5)) # графічне подання результатів обчислення експерименту
```

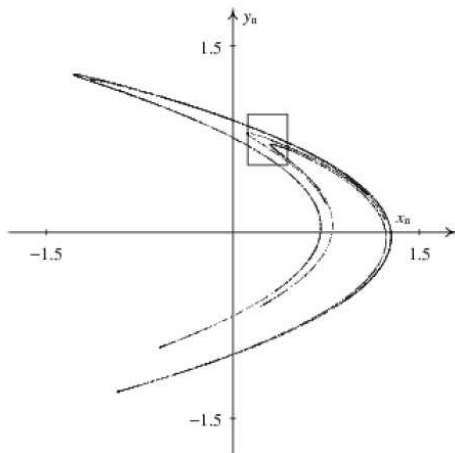


Рис. В.2. Карта Хенона для рівняння  $y = 1 + 0.3 \cdot y - 1.4 \cdot x^2$

Дотримання архітектури MVC дозволяє збільшити невелику область з центром в  $(0.31, 0.68)$  без повторного розрахунку моделі:

```
sage: b=point(a,pointsize=5)
b.set_axes_range(0.12,0.44,0.54,0.95)
```

show(b)

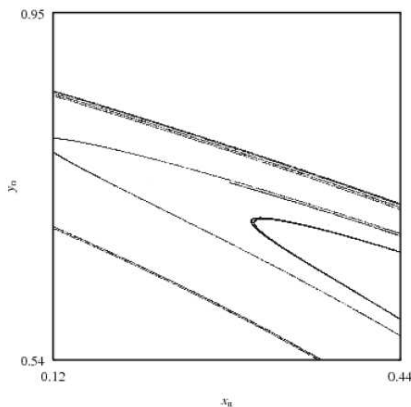


Рис. В.3. Самоподібність карт Хенона

Графіки на рис. В.2–3 є фракталами, про що свідчить його самоподібність (існування частин, які повторюються, при довільному збільшенні будь-якої частини зображення). Рис. В.3 отримано за допомогою збільшення частини зображення, яке виділене квадратом на рис. В.2.

Функція `chaosgame(point, p0, b, n)` реалізує так звану гру хаосу: спочатку розраховується точка  $p_0 = [x_0, y_0]$ , що належить списку `point = ([x1, y1], ..., [xm, ym])`, яка обирається у випадковий спосіб. Наступна точка розраховується із відрізка, що з'єднує попередню точку з випадково обраною на відстані, рівній довжині цього відрізка, помноженій на  $b$ . Процедура повторюється  $n$  разів.

Виконавши команду:

```
sage: c1=chaosgame([[0, 0], [1, 0], [0.5, 0.866025]],  
[0.1, 0.1], 1/2, 30000)
```

отримаємо фрактальну поверхню – серветку Серпінського (рис. В.4). На рис. В.5 зображено дивний атрактор в 5 точках з вершиною в центрі площини та половинним коефіцієнт стиснення.

```
sage: c2=chaosgame([[0, 0], [0, 1], [1, 0], [1, 1],  
[0.5, 0.5]], [0.5, 0.5], 1/3, 30000)  
show(point(c1))  
show(point(c2))
```

Функція `ifs(prob, mat, point, p0, n)` реалізує метод системи ітеративних функцій. Параметрами функції є перелік ймовірних значень `prob` із списку  $m$  матриць `mat`, списку  $m$  точок

point, одна з яких є початковою  $p_0$  та кількістю ітерацій  $n$ , яку треба виконати.

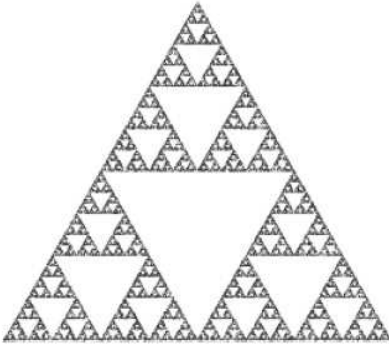


Рис. В.4. Серветка Серпінського

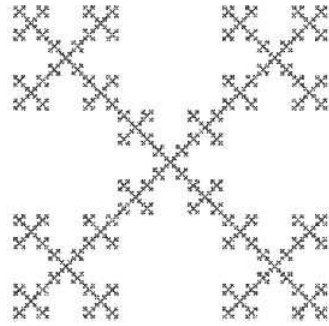


Рис. В.5. Дивний атрактор

Застосувавши наступні команди, отримаємо фрактальне дерево (рис. В.6):

```
sage: m1=matrix([[ -0.550, -0.179], \
[-0.179, 0.550]])
p1=matrix([[ -0.438], [0.382]])
m2=matrix([[ -0.246, 0.193], \
[0.275, 0.365]])
p2=matrix([[ -0.379], [0.538]])
m3=matrix([[0.006, -0.014], \
[-0.147, -0.459]])
p3=matrix([[ -0.283], [0.490]])
c=ifs([1,2,3],[m1,m2,m3],[p1,p2,p3], \
[-0.3,0.2],10000);
show(point(c))
```

Для побудови спочатку ми використовували в якості початкової точки початок координат та починали з малого значення  $n$ , а для збільшення рівня деталізації використали в якості початкової точки  $(-0.3, 0.2)$ , що є частиною фракталу, і брали більшу кількість ітерацій.

Засобами функції `ifs` можна отримати дивну папороть Баренслі:

```
sage: a1=matrix([[0.85,0.04],[ -0.04,0.85]])
a2=matrix([[0.2, -0.26],[0.23,0.22]])
a3=matrix([[ -0.15,0.28],[0.26,0.24]])
a4= matrix([[0,0],[0,0.16]])
p1=matrix([[0],[1.6]])
p2=matrix([[0],[1.6]])
```

```

p3=matrix([[0],[0.44]])
p4=matrix([[0],[0]])
sage: c=ifs([85,92,99,100],[a1,a2,a3,a4],\
[p1,p1,p3,p4],[5,0],50000);
show(point(c,rgbcolor=(0,1,0)))

```

Функція `staircase (fun, initial, n)` розраховує множину точок для побудови ітераційної діаграми послідовності з  $n$  елементів, що задається рекурентним співвідношенням  $fun$  ( $y_{n+1}=F(y_n)$ ) з початковим значенням `initial`.



Рис. В.6. Фрактальне дерево, що утворюється системою з трьох ітераційних функцій

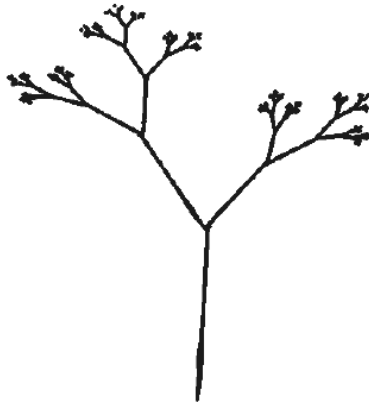


Рис. В.7. Фрактальний об'єкт «дивний папороть Баренслі», що утворюються системою з чотирьох функцій

Для побудови ітераційної діаграми для послідовності  $y_{n+1}=y_n^2-0.2$  з початковим значенням 1.1 (рис. 8) слід виконати команду:

```

sage: var('y')
p=staircase(y^2-0.2,1.1,10)
show(point(p))

```

Функція `orbits (f, y0, n1, n2, domain)` розраховує траєкторію сімейства одновимірних дискретних динамічних систем  $f$  з одним параметром  $x$ ,  $y_0$  – початкове значення,  $n1$  – кількість ітерацій, що будуть ігноруватися,  $n2$  – кількість ітерацій, які будуть зображені на рисунку, `domain` – список, який починається з  $x$  і визначає початкове значення для перелічених в ньому параметрів, останнє значення даного списку – приріст параметра.



Даний тип діаграм використовується при вивченні біфуркації  
одновимірних дискретних систем.

Побудуємо траєкторію для квадратичного відображення  
 $y_{n+1} = y_n^2 + x$ , де  $x$  – параметр (рис. В.9-10):

```
sage: var('y')
      a=orbits(x+y^2,0,100,400,\
              [x,y,0,-1.53,-0.001]);
      show(b)
```

Дотримання архітектури MVC дозволяє збільшити невелику  
область з центром  $(-1.25, -1.25)$  без повторного розрахунку моделі.  
Для цього необхідно виконати наступні команди:

```
sage: b.set_axes_range(-1.5,-1,0.8,1.6)
      b=point(a,pointsize=3)
      show(b)
```

В результаті отримуємо рис. В. 10, що демонструє самоподі-  
бність, а отже й фрактальну природу досліджуваного об'єкту.

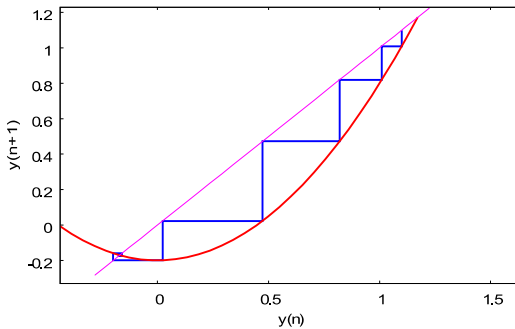


Рис. В.8. Графічне подання ітераційної діаграми  
для послідовності  $y_{n+1} = y_n^2 - 0.2$

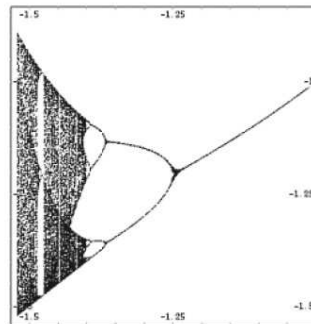
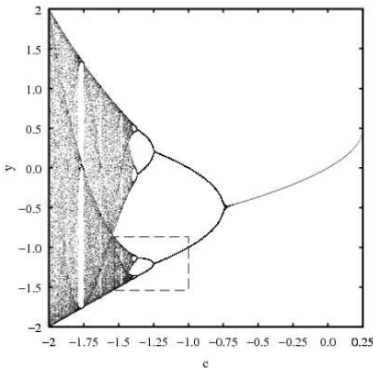


Рис. В.9. Траекторія руху динамічної системи  $y_{n+1}=y_n^2+x$

Рис. В.10. Збільшена область динамічної системи  $y_{n+1}=y_n^2+x$ , в області з центром  $(-1.25, -1.25)$

## Література

1. Гулд Х. Компьютерное моделирование в физике : часть 1 / Х. Гулд, Я. Тобочник. – М. : Мир, 1990. – 350 с.
2. Гулд Х. Компьютерное моделирование в физике : часть 2 / Х. Гулд, Я. Тобочник. – М. : Мир, 1990. – 350 с.
3. **Зализняк В. Е. Основы научных вычислений. Введение в численные методы для физиков : учебное пособие для студентов естественно-научных и технических специальностей высших учебных заведений. / В. Е. Зализняк- М. : Эдиториал УРСС, 2002.- 296 с.**
4. **Мантенья Р. Н. Введение в эконофизику. Корреляции и сложность в финансах / Мантенья Р. Н., Стенли Х. Ю. - М., 2007. - 188 с.**
5. Мандельброт Б. Фрактальная геометрия природы / Б. Мандельброт. – М. : Институт компьютерных исследований, 2002. – 656 с.
6. Петерс Э. Фрактальный анализ финансовых рынков : применение теории хаоса в инвестициях и экономике / Э. Петерс. – М. : Интернет-трейдинг, 2004. – 304 с.
7. Синергетичні та еконофізичні методи дослідження динамічних та структурних характеристик економічних систем : монографія / В. М. Соловійов, В. Д. Дербенцев, О. А. Сердюк, О. Д. Шарапов. – Черкаси : Брама-Україна, 2010. – 287 с.
8. Соловійов В. М. Математична економіка : навчально-методичний посібник для самостійного вивчення дисципліни / В. М. Соловійов. – Черкаси, 2008. – 137 с.
9. Сорнетте Д. Как предсказывать крахи финансовых рынков. Критические события в комплексных финансовых системах / Д. Сорнетте. – М., 2003. – 394 с.
10. Шокалюк С. В. Основы работы в Sage / С. В. Шокалюк ; за ред. академіка АПН України М. І. Жалдака. – К. : НПУ імені М. П. Драгоманова, 2008. – 64 с.
11. Eckmann J.-P. Recurrence Plots of Dynamical Systems / J.-P. Eckmann, S. Oliffson Kamphorst and D. Ruelle // Europhysics Letters. – 1987. – Volume 4, Number 9. – P. 973–977.
12. Mantegna R. N. An Introduction to Econophysics / Mantegna R. N., Stanley H. E. – Cambridge : Cambridge

- University Press, 2000. – 144 p.
13. Mantegna R. N. Information and hierarchical structure in financial markets // *Computer Physics Communications*. – 1999. – Vol. 121–122. – P. 153–156.
  14. Mehta M. L. *Random matrices* / M. L. Mehta. – New York : Academic Press, 1991. – 688 p.
  15. Random matrix approach to cross correlations in financial data / Plerou V., Gopikrishnan P., Rosenow B., Amaral L. A. N., Guhr T., Stanley H. E. // *Phys. Rev. E*. – 2002. – V. 65, №12. – P. 356-373.
  16. Zbilut J. P. Recurrence quantification based Liapunov exponent for monitoring divergence in experimental data / Zbilut J. P, Zaldivar-Comenges J.- M., Stozzi F. // *Phys. Lett.* – 2002. – V. 297. – P. 173–181.
  17. Stein W. *Sage Programming Guide* / Stein W., Joyner D. – 2008. – 86 p.

Соловйов Володимир Миколайович  
Соловйова Вікторія Володимирівна  
Хараджян Наталя Анатоліївна

# Моделювання складних економічних систем

*Навчальний посібник*

Підп. до друку 24.11.2010  
Папір офсетний №1  
Ум. друк. арк. 6,9

Формат 80×84 1/16  
Зам. №2-1110  
Тираж 300 прим.

Жовтнева районна друкарня  
50014, м. Кривий Ріг, вул. Електрична, 5  
Тел. (0564) 407-29-02  
E-mail: [nata\\_leonova@mail.ru](mailto:nata_leonova@mail.ru)