

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ  
УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА

Економічний факультет

ІНСТИТУТ МЕНЕДЖМЕНТУ  
ТА ФІНАНСІВ

при  
Київському національному  
університеті імені Тараса Шевченка

Навчально-методичний комплекс

з курсів

**„Прогнозування” та „Фінансове прогнозування”**

для студентів економічних спеціальностей  
денної, очно-заочної та заочної форм навчання

Київ – 2006

**Ставицький А.В.**

Навчально-методичний комплекс з курсів „Прогнозування” та „Фінансове прогнозування”. – К., 2006. – 107 с.

Розглянуто та схвалено на засіданні кафедри економічної кібернетики, протокол № 11 від 14 лютого 2006 р.

Схвалено Вченою радою економічного факультету Київського національного університету імені Тараса Шевченка, протокол № 7 від 28 лютого 2006 року.

Рецензенти: **О.І.Черняк**, д.е.н., професор, завідувач кафедри економічної кібернетики Київського національного університету імені Тараса Шевченка;

**О.В.Комашко**, к.ф.-м.н., доцент кафедри економічної кібернетики Київського національного університету імені Тараса Шевченка.

## Зміст

Загальні положення .....	4
Анотація курсу .....	4
Програма курсу .....	5
Тематичний план дисципліни .....	6
Теми лекцій .....	7
Теми лабораторних робіт .....	9
Розділ 1. Система і принципи економічного прогнозування .....	10
Теоретична частина .....	10
Практична частина .....	11
Лабораторна робота .....	19
Розділ 2. Прогнозування за допомогою часових рядів .....	20
Теоретична частина .....	20
Практична частина .....	26
Лабораторна робота .....	30
Розділ 3. Методи згладжування часових рядів .....	31
Теоретична частина .....	31
Практична частина .....	34
Лабораторна робота .....	37
Розділ 4. Виділення трендового та сезонного компонентів .....	38
Теоретична частина .....	38
Приклад .....	39
Практична частина .....	40
Лабораторна робота .....	47
Розділ 5. ARIMA–моделі .....	49
Теоретична частина .....	49
Практична частина .....	59
Лабораторна робота .....	64
Розділ 6. Моделі зі змінною дисперсією .....	66
Теоретична частина .....	66
Практична частина .....	69
Лабораторна робота .....	70
Розділ 7. VAR-моделі .....	72
Теоретична частина .....	72
Практична частина .....	76
Лабораторна робота .....	83
Розділ 8. Сучасне прогнозування .....	84
Теоретична частина .....	84
Лабораторна робота .....	88
Програмування у EViews .....	89
Завдання для самостійної роботи з курсу .....	98
Контрольні роботи з курсу .....	100
Контрольна робота для стаціонарної форми навчання .....	100
Контрольна робота для заочної та очно-заочної форм навчання .....	101
Питання на залік з курсу .....	104
Література .....	106

## ЗАГАЛЬНІ ПОЛОЖЕННЯ

### *Анотація курсу*

В рамках курсів „Прогнозування” та „Фінансове прогнозування” поглиблюються та розширюються предметні області і методологічні засоби курсів „Економетрика”, „Прикладна економетрика”, „Статистика”. Основна увага приділяється використанню засобів комп’ютерної техніки для здійснення аналізу економічних змінних та їх прогнозування.

Для правильного використання методів прогнозування в емпіричних дослідженнях необхідно його засвоїти та адекватно застосовувати на практиці. Аналіз часових рядів є складною темою, і тому ця робота має за мету допомогти студентам застосувати набуті знання на реальних економічних даних.

Основною метою курсу стає можливість студента

- 1) зробити прогноз досліджуваного процесу на певний час у майбутнє;
- 2) визначити, як потрібно впливати на об’єкт, щоб одержати задані параметри його функціонування;
- 3) визначити, чи існують взаємозв'язки між обраними параметрами процесу, і який їх вид;
- 4) визначити, чи існує взаємозв'язок між деякими процесами.

**Мета вивчення дисципліни** – ознайомлення студентів з методами прогнозування економічної інформації на основі аналізу статистичних даних. Знання, здобуті студентами під час вивчення курсів „Прогнозування” та „Фінансове прогнозування”, широко застосовуються в менеджменті, маркетингу, мікро- та макроекономіці, фінансовій справі, при виконанні творчих індивідуальних завдань, курсових робіт та написанні дипломних проектів.

**Навчальна задача курсу** полягає у розвитку аналітичних та практичних здібностей студентів у прогнозуванні економічної інформації.

**Зв'язок з іншими навчальними дисциплінами.** Базовими для курсів „Прогнозування” та „Фінансове прогнозування” є дисципліни економічного циклу, такі як „Економічна теорія”, „Мікроекономіка”, „Макроекономіка”, „Фінанси”. Математичною основою курсу є дисципліна „Теорія ймовірностей та математична статистика” та „Економетрика”.

**Методи проведення занять.** Поєднання лекційних занять з роботою в комп’ютерному класі. **Лекції** націлені забезпечити теоретичне підґрунтя курсу, розкрити зміст основних методів аналізу економічної інформації. **Лабораторні заняття** передбачають практичне застосування методів економетричного аналізу над економічними даними.

**Форма контролю.** Оцінка виставляється за результатами виконання спеціального проекту протягом вивчення курсу, написання заключного тесту та роботи на семінарських заняттях.

## ***Програма курсу***

### **Тема 1. Система і принципи економічного прогнозування**

Предмет курсу. Прогноз. Економічний прогноз. Прийом прогнозування. Метод прогнозування. Методика прогнозування. Прогнозна модель. Система прогнозування. Ефективність прогнозування. Функції економічного прогнозу. Класифікація прогнозів. Послідовність операцій розробки прогнозу.

### **Тема 2. Прогнозування за допомогою часових рядів**

Основні визначення. Порядок аналізу часових рядів. Адитивна та мультиплікативна моделі часових рядів. Міри точності прогнозів. Стаціонарність часових рядів. Функція автокореляції. Функція правдоподібності. Максимізація функції правдоподібності. Перевірка гіпотез за допомогою функції правдоподібності. Числові характеристики часових рядів. Корелограма. Стабільність моделі. Метод поворотних точок.

### **Тема 3. Методи згладжування часових рядів**

Класичні підходи (метод усереднення, подвійне усереднення). Методи експоненціального згладжування: звичайне, подвійне, потрійне. Адаптивне згладжування. Несезонна модель Холта–Вінтерса. Адитивна модель із визначенням сезонних коливань. Адитивна модель Холта–Вінтерса. Мультиплікативна модель Холта–Вінтерса. Фільтр Ходріка–Прескотта.

### **Тема 4. Виділення трендового та сезонного компонентів**

Часові ряди з трендом. Функціональні види тренду. Виділення сезонних коливань. Фіктивні змінні. Моделювання сезонності. Прогнозування сезонності в макроекономічних дослідженнях.

### **Тема 5. ARIMA-процеси**

Поняття «білого шуму». Лаговий оператор.  $MA(q)$ -процес. Процес авторегресії. Перетворення  $MA$ -процесів.  $ARMA$ -процес. Прогнозування на основі  $ARMA$ -моделей. Оцінювання невідомих коефіцієнтів.  $ARIMA$ -процеси. “Випадкове блукання”. Тестування на наявність одиничного кореня. Імпульсний аналіз. Аналіз часових рядів Бокса–Дженкінса: ідентифікація моделі, оцінка моделі, діагностика моделі. Аналіз залишків моделі, порівняння моделей, числові критерії адекватності моделей.

## Тема 6. Моделі зі змінною дисперсією

Приклади нелінійних процесів: нелінійні процеси авторегресії, порогові моделі, білінійні моделі, гармонічні процеси. Моделювання фінансових ринків. Моделі зі змінною дисперсією: ARCH, GARCH, TARARCH, EGARCH, ARCH-M та інші. Оцінювання моделей зі змінною дисперсією. Моделювання часових рядів при зміні економічної ситуації. Економічний аналіз на основі моделей зі зміною економічних ситуацій.

## Тема 7. VAR-моделі

Означення VAR-моделі. Оцінка стаціонарних VAR-моделей. Прогнозування на основі VAR-моделей. Структурний аналіз на основі VAR-моделей. Причинність за Гренджером. Імпульсний аналіз. Різновиди VAR-моделей. Коінтеграція. Тестування на наявність коінтеграції.

## Тема 8. Сучасні проблеми прогнозування

Комбінування прогнозів. Аналіз нових задач при прогнозуванні часових рядів. Узагальнення теорії прогнозування. Нові шляхи прогнозування. Принципи вибору моделей для прогнозування. Експертне прогнозування. Методи вибіркового дослідження при прогнозуванні часових рядів. Прогнозування на основі нейронних мереж. Проблема дезагрегування часових рядів.

## Тематичний план дисципліни

Назви розділів та тем	Кількість навчальних годин			
	Денна ф.н./оч.-заоч./заоч.ф.н.			
	Всього	Лекції	Лаборат. Заняття	Самост. робота
Система і принципи економічного прогнозування	6/6/6	2/1/1	2/-/-	2/5/5
Прогнозування за допомогою часових рядів	6/6/6	2/1/1	2/1/1	2/4/4
Методи згладжування часових рядів	6/6/6	2/2/1	2/2/1	2/2/4
Виділення трендового та сезонного компонентів	6/6/6	2/2/1	2/1/1	2/3/4
ARIMA-моделі	8/8/8	4/2/-	2/1/1	2/5/7
Моделі зі змінною дисперсією	10/10/10	2/2/-	4/2/-	4/6/10
VAR-моделі	6/6/6	2/1/-	2/-/-	2/5/6
Сучасні проблеми прогнозування	6/6/6	2/1/-	2/1/-	2/4/6
	<b>54/54/54</b>	<b>18/12/4</b>	<b>18/8/4</b>	<b>18/34/46</b>

## **Тема лекцій**

### **Лекція 1. Система і принципи економічного прогнозування**

1. Предмет курсу.
2. Функції економічного прогнозу.
3. Класифікація прогнозів.
4. Послідовність операцій розробки прогнозу.

*Література: [2, с. 5-19], [4]*

### **Лекція 2. Прогнозування за допомогою часових рядів**

1. Основні визначення.
2. Порядок аналізу часових рядів.
3. Адитивна та мультиплікативна моделі часових рядів.
4. Міри точності прогнозів.
5. Стаціонарність часових рядів.
6. Числові характеристики часових рядів.

*Література: [10, с.7-15], [8, с. 147-165]*

### **Лекція 3. Методи згладжування часових рядів**

1. Метод усереднення. Метод подвійного усереднення.
2. Методи експоненціального згладжування.
  - а. Звичайне експоненціальне згладжування.
  - б. Подвійне експоненціальне згладжування.
  - с. Потрійне експоненціальне згладжування.
  - д. Адаптивне згладжування.
3. Несезонна модель Холта–Вінтерса.
4. Адитивна модель Холта–Вінтерса.
5. Мультиплікативна модель Холта–Вінтерса.
6. Фільтр Ходріка–Прескотта.

*Література: [10, с.17-23, 49-50], [8, с. 166-173, 200-203], [5], [6]*

### **Лекція 4. Виділення трендового та сезонного компонентів**

1. Часові ряди з трендом.
2. Функціональні види тренду.
3. Виділення сезонних коливань. Фіктивні змінні.
4. Прогнозування сезонності в макроекономічних дослідженнях.

*Література: [10, с.48-49, 59-60], [8, 197-200, 210-213], [11], [5], [6]*

### **Лекція 5. ARMA-процеси**

1. Поняття «білого шуму».
2. Лаговий оператор.
3.  $MA(q)$ -процес.
4. Процес авторегресії.
5. Перетворення  $MA$ -процесів.

6. ARMA–процес.
  7. Прогнозування на основі ARMA-моделей.
  8. Оцінювання невідомих коефіцієнтів ARMA-моделей.
- Література: [10, с. 27-41], [8, с. 179-191], [6]*

### Лекція 6. ARIMA-процеси

1. ARIMA-процеси.
  2. “Випадкове блукання”.
  3. Тестування на наявність одиничного кореня.
  4. Імпульсний аналіз.
  5. Аналіз часових рядів Бокса-Дженкінса.
    - a. Ідентифікація моделі.
    - b. Оцінка моделі.
    - c. Діагностика моделі. Числові критерії адекватності моделей.
  6. Аналіз залишків моделі
- Література: [10, с. 41-47], [8, с. 191-196], [7, с. 239-255], [6]*

### Лекція 7. Моделі зі змінною дисперсією

1. Приклади нелінійних процесів.
  2. Моделі зі змінною дисперсією.
  3. Оцінювання моделей зі змінною дисперсією.
  4. Моделювання часових рядів при зміні економічної ситуації.
  5. Економічний аналіз на основі моделей зі зміною економічних ситуацій.
- Література: [10, с. 60-74], [8, 214-236], [7, с. 275-278], [6]*

### Лекція 8. VAR-моделі

1. Означення VAR–моделі. Різновиди VAR-моделей.
  2. Оцінка стаціонарних VAR-моделей.
  3. Прогнозування на основі VAR-моделей.
  4. Структурний аналіз на основі VAR-моделей.
  5. Причинність за Гренджером.
  6. Імпульсний аналіз.
  7. Коінтеграція. Тестування на наявність коінтеграції.
- Література: [10, с. 75-81], [8, 237-244], [7, с.239-249], [6]*

### Лекція 9. Сучасні проблеми прогнозування

1. Комбінування прогнозів.
  2. Узагальнення теорії прогнозування. Нові шляхи прогнозування.
  3. Принципи вибору моделей для прогнозування.
  4. Експертне прогнозування.
  5. Проблема дезагрегування часових рядів.
- Література: [10, с. 24-26, 82-87], [8, с. 245-260, 176-178], [9], [6]*



## **Теми лабораторних робіт**

Лабораторна робота 1. Робота з EViews.

*Література:* [15], [18]

Лабораторна робота 2. Метод поворотних значень.

*Література:* [12], [15], [18]

Лабораторна робота 3. Методи згладжування економічної інформації.

*Література:* [12], [15], [18]

Лабораторна робота 4. Виділення трендового та сезонного компонентів з  
часового ряду

*Література:* [12], [15], [18]

Лабораторна робота 5. Ідентифікація ARIMA-моделей.

*Література:* [6], [12], [15], [18]

Лабораторна робота 6. Прогнозування за допомогою ARIMA-моделей.

*Література:* [6], [12], [15], [18]

Лабораторна робота 7. Прогнозування за допомогою процесів зі змінною  
дисперсією.

*Література:* [12], [15], [18]

Лабораторна робота 8. Оцінка VAR-моделей

*Література:* [6], [12], [15], [18]

Лабораторна робота 9. Методи дезагрегування часових рядів.

*Література:* [10], [12], [15], [18]

# РОЗДІЛ 1. СИСТЕМА І ПРИНЦИПИ ЕКОНОМІЧНОГО ПРОГНОЗУВАННЯ

## *Теоретична частина*

### Основні визначення

**Предметом курсу** є сукупність методів, засобів та способів розробки фінансових прогнозів.

Під **прогнозом** розуміється науково-обґрунтоване судження про можливі стани об'єкта в майбутньому, про альтернативні шляхи і терміни його здійснення. Процес розробки прогнозів називається **прогнозуванням**, а наукові дисципліни про закономірності розробки прогнозів - **прогностикою**.

**Економічний прогноз** – це висловлення (судження), що обмежене одним або декількома економічними аспектами прогнозованих подій.

У економічному прогнозуванні виділяються такі основні поняття.

**Прийом прогнозування** – одна або декілька математичних або логічних операцій, спрямованих на одержання конкретного результату в процесі розробки прогнозу.

**Метод прогнозування** – сукупність прийомів, спрямованих на розробку прогнозу.

**Методика прогнозування** – сукупність методів і правил розробки прогнозів.

**Об'єкт прогнозування** – процеси, явища і події, на які спрямована діяльність суб'єкта прогнозування.

**Прогнозна модель** – модель об'єкта прогнозування, дослідження якої дозволяє одержати інформацію про можливі стани об'єкта в майбутньому.

**Система прогнозування** – сукупність методик, технічних засобів, призначених для прогнозування складних явищ або процесів.

**Ефективність прогнозування** визначає необхідність перевищення економічного ефекту від використання прогнозу над витратами по його розробці.

### Функції економічного прогнозу

До **основних функцій** економічного прогнозу відносяться:

- науковий аналіз економічних, соціальних і науково-технічних процесів і тенденцій, об'єктивних зв'язків цих процесів у конкретних умовах у визначеному періоді;
- оцінка об'єкта прогнозування;
- виявлення альтернатив розвитку процесів;
- накопичення економічної інформації для прийняття оптимального рішення.

## Класифікація прогнозів

1. За видом. Загальний прогноз – відповідає на питання про можливість або неможливість настання деякої події; якісний прогноз – передбачає опис майбутньої ситуації; кількісний прогноз – передбачає визначення числових характеристик майбутньої ситуації на основі точкових або інтервальних оцінок.

2. За періодом прогнозування. Оперативні (поточні), коротко-, середньо-, довго- і дальньострокові прогнози. Оперативні прогнози (до 1 місяця) містять, як правило кількісні оцінки, короткострокові (від 1 місяця до року) – загальні кількісні оцінки, середньострокові (від 1 року до 5 років) – кількісно-якісні, довгострокові (від 5 до 15 років) – якісно-кількісні, дальньострокові (понад 15 років) – загальні якісні оцінки.

3. За критерієм масштабності об'єкта. Сублокальні (число змінних від 1 до 3); локальні (від 4 до 14); субглобальні (число змінних від 15 до 35); глобальні (число змінних від 36 до 100); суперглобальні (число змінних понад 100).

4. За критерієм складності. Надпростий, простий, складний, надскладний. Ці прогнози різняться наявністю взаємозалежних змінних у їхньому описі: у надпростому прогнозі відсутні істотні взаємозв'язки між змінними; у простому досліджуються парні взаємозв'язки між змінними; у надскладному враховуються взаємозв'язки між усіма перемінними,

5. За ступенем детермінованості об'єкта. Детерміновані, стохастичні (у яких враховується випадкова складова), змішані.

## Послідовність операцій розробки прогнозу

1. **Передпрогнозна орієнтація** (програма дослідження) включає: уточнення завдання на прогноз (масштаб, об'єкт, попередження і т.д.); формулювання цілей і задач прогнозу, визначення структури й організації дослідження.

2. **Прогнозна ретроспекція** – дослідження історії розвитку об'єкта прогнозування з метою одержання його систематизованого опису.

3. **Прогнозний діагноз** – дослідження систематизованого опису об'єкта прогнозування з метою виявлення тенденції його розвитку та вибору або розробки моделей і методів прогнозування.

4. **Прогнозна перспекція** – розробка прогнозу за результатами прогнозного діагнозу.

5. **Верифікація прогнозу** – оцінка достовірності і точності, а також обґрунтованості прогнозу.

6. **Коригування прогнозу** – уточнення прогнозу на підставі його верифікації та додаткових даних.

7. **Синтез прогнозів** – розробка системного прогнозу.

## Практична частина

Використання методів прогнозування важко уявити без відповідних комп'ютерних програм. З одного боку хочеться мати найпотужнішу систему, яка може обробляти довільну інформацію, в якій запрограмовані сотні

різноманітних методів аналізу. З іншого боку, складність роботи з системою має бути мінімальною. Ніяким чином не заперечуючи право кожного на вільний вибір статистичного пакету для роботи у підручнику представлена одна з найкращих та потужніших систем для аналізу та прогнозування часових рядів Econometrics Views.

Майже кожен, хто знайомий з цією системою відмічає її безперечні переваги: простота у роботі, графічне оформлення результатів, набір багатьох методів аналізу, можливість створювати свої методи. Цих ознак виявилось достатньо, щоб пакет EViews став провідним у використанні у аналітичних відділах міністерств, банків, фірм, організацій.

Пакет EViews реалізований як стандартна програма під операційну систему Windows, що робить її легкою у використанні навіть для непрофесіоналів. Одним з недоліків системи залишається її англійський інтерфейс та вимога створювати та обробляти всі файли, об'єкти лише англійською мовою.

### Створення робочої книги

Робота з програмою починається зі створення файлу, в якому будуть зберігатися дані, побудовані моделі, системи тощо. Для створення нового файлу слід скористатися меню системи **File→New→WorkFile**. У вікні необхідно вказати періодичність даних, а також початковий та кінцевий періоди аналізу. Формат дати є досить простим. Для річних даних слід вказати лише номер початкового та кінцевого років, наприклад, 1998. Для піврічних, квартальних та місячних даних слід вказати рік і через двокрапку номер півріччя, кварталу чи місяця відповідно, наприклад, як це показано на рис. 1. Для добових даних вказується повністю дата, наприклад, запис 19:10:2005 відповідає даті 19 жовтня 2005 року. Нарешті, при введенні послідовних даних (*undated* або *irregular*) слід вказати індекс першого та останнього елементів, наприклад, 1 та 24 відповідно.

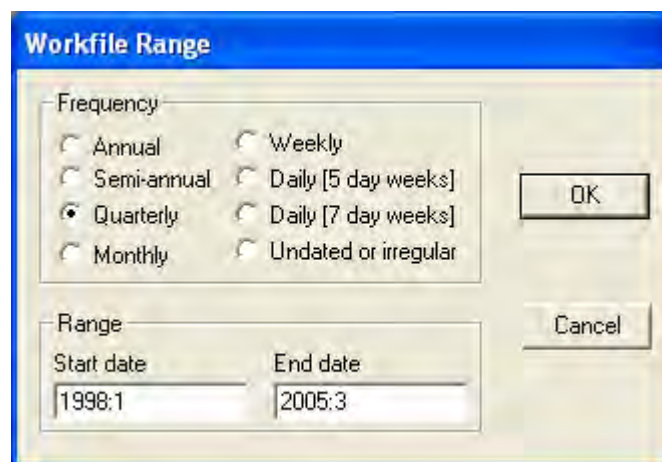


Рис. 1. Вказування періодичності та діапазону даних

При закритті вікна створюється новий файл, вигляд якого показано на рис. 2.

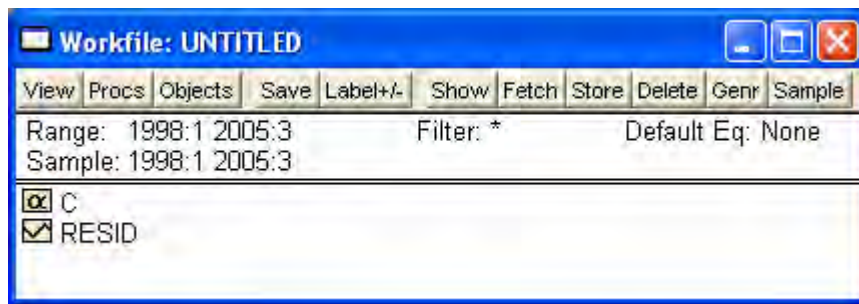




Рис. 2. Вікно новоствореного файлу

Верхня частина вікна являє собою меню. Нижче вказаний діапазон (*Range*) даних. Це означає, що у файлі міститимуться лише дані з 1 кварталу 1998 року по 3 квартал 2005 року. Для зміни цього показника необхідно двічі клацнути мишею на полі *Range*, або вибрати меню **Procs**→**Change Workfile range....**

Наступний елемент (*Sample*) вказує на вибірку, яка буде враховуватися при проведенні різноманітних операцій і оцінці моделей. Ця вибірка не може бути більшою за діапазон файлу. Зміна цього параметру здійснюється за допомогою подвійного натискання мишею на полі *Sample*, або за допомогою меню **Procs**→**Sample....** При цьому можна не тільки вручну ввести необхідні початкову та кінцеві дати, а й вказати умову відбору, наприклад, до вибірки можна включити лише ті елементи, для яких значення змінної більше певного рівня ( $a > 1876$ ).

Нижче у вікні показані існуючі об'єкти. EViews автоматично створює 2 системні об'єкти. Перший з них називається *C* і містить нульові значення. Він виконує роль константи при побудові моделі. Всі аналогічні об'єкти позначатимуться за допомогою символу . Змінна *RESID* міститиме залишки при побудові моделей. Кожного разу при оцінюванні нової моделі ця змінна буде наповнена новими значеннями. Всі змінні в EViews позначаються за допомогою символу .

Фільтрацію відображених об'єктів можна здійснити за допомогою елемента *Filter* (подвійне натиснення мишею на полі *Filter*) або за допомогою меню **View**→**Display Filter...** (рис.3).

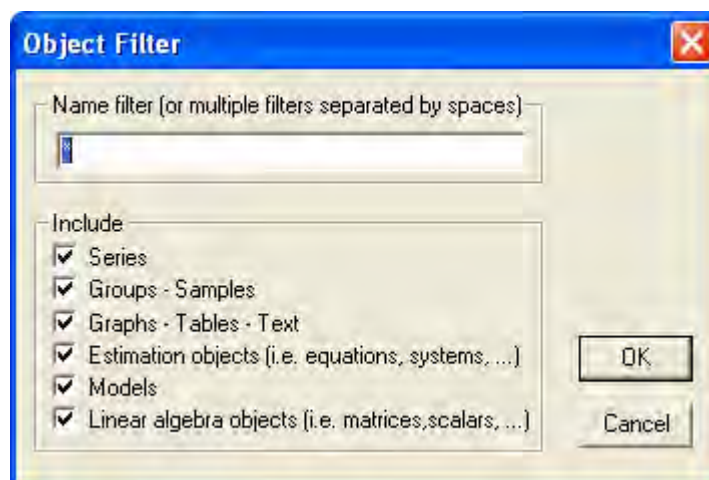


Рис. 3. Застосування фільтрації об'єктів

За допомогою фільтру можна показувати лише об'єкти визначеного типу, наприклад, лише часові ряди (*Series*), групи вибірок (*Group – Samples*), побудовані регресії (*Estimation objects*), системи рівнянь (*Models*) тощо.

Крім цього фільтрація дозволяє виводити лише ті об'єкти, назви яких відповідають певній масці. Наприклад, якщо у верхньому рядку вікна написати „C\*“, то будуть виведені всі об'єкти, що починаються з літери „C“. Така можливість є зручною при роботі з великою кількістю змінних.

Для запису створеного файлу на диск потрібно скористатися меню **File→Save as...**, вказавши необхідну папку та назву файлу (рис.4).



Рис. 4. Збереження файлу на диску

Слід пам'ятати, що EViews не сприймає кириличні символи, тому назву файлу та назви папок потрібно називати лише латинськими літерами та цифрами.

Відкриття файлу здійснюється за допомогою команди **File→Open Workfile....**

### Створення змінних та імпорт даних

Для введення даних до програми існує декілька можливостей.

Перша полягає у створенні об'єкта-змінної та ручного набору необхідної інформації. Для реалізації цієї можливості слід вибрати меню **Objects→New Object....** Серед запропонованих варіантів обирається тип *Series* та вказується назва змінної (рис. 5).

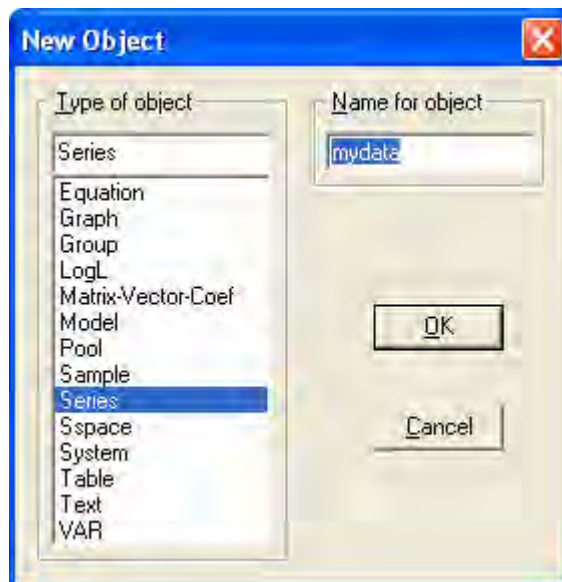


Рис. 5. Створення нової змінної

При цьому до списку об'єктів робочої книги додається змінна *mydata*. Для перегляду її значень двічі натискаємо на об'єкті. Всі значення змінної ще є невизначеними. Щоб розпочати введення даних слід переключитися до режиму редагування за допомогою кнопки меню **Edit+/-**. Після цього інформація вводиться до чарунок листа (рис. 6). Аналогічним чином проводиться і редагування змінної. Іноді при цьому корисно представити інформацію не у стовпчику, а у більш структурованому вигляді. Перемикач **Wide+/-** дозволяє перемикатися до перегляду даних поквартально.

Series: MYDATA Workfile: MODEL\_1

View Procs Objects Print Name Freeze Transform Edit+/- Smpl+/- Label+/- Wide+/-

1281 MYDATA

Last updated: 01/12/06 - 14:30

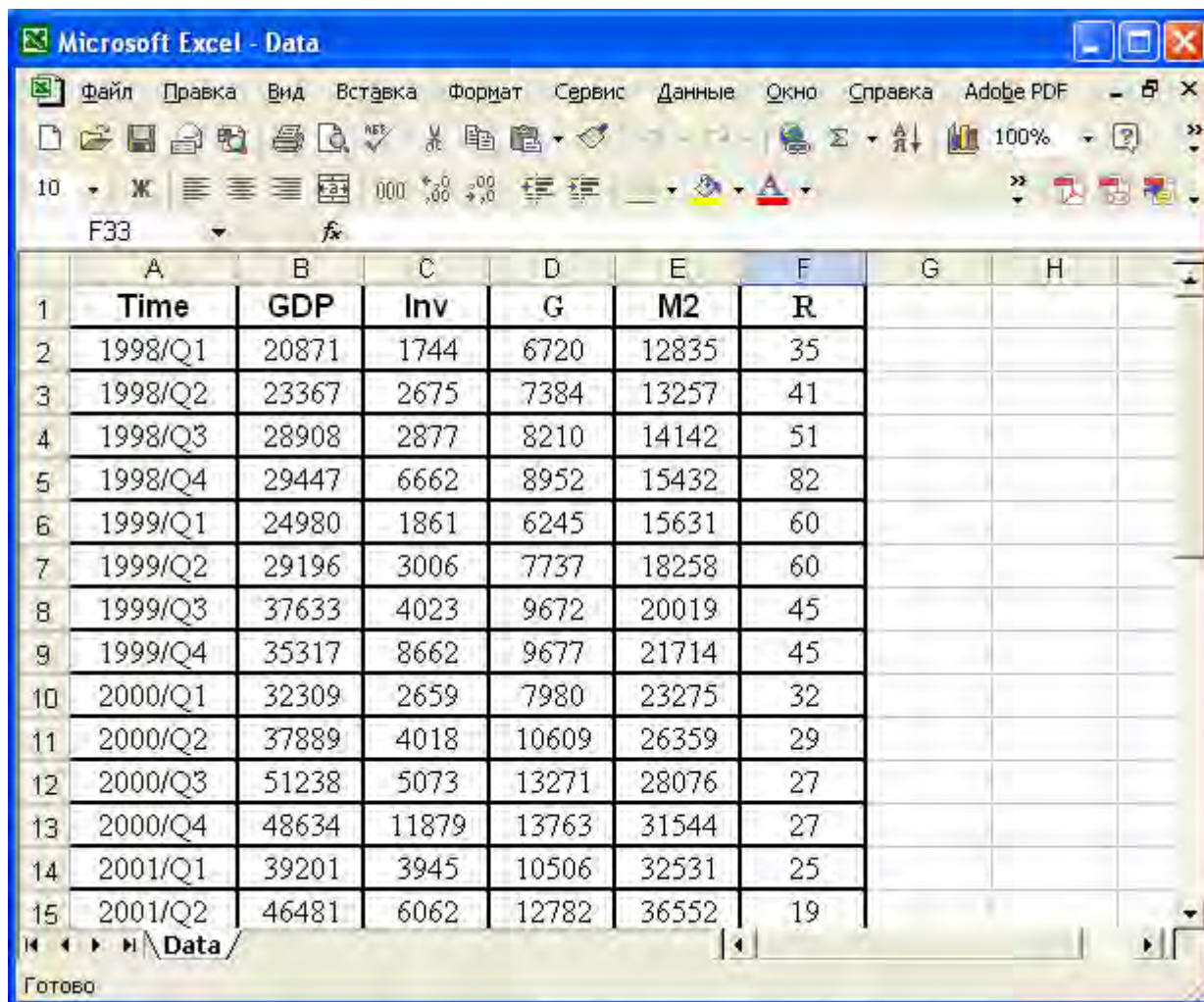
1998:1	1123.000			
1998:2	1256.000			
1998:3	1281			
1998:4	NA			
1999:1	NA			
1999:2	NA			
1999:3	NA			
1999:4	NA			
2000:1	NA			
2000:2				

Рис. 6. Введення та редагування змінної.

Друга можливість для створення нових змінних полягає у імпортуванні їх з інших джерел. Найчастіше таким джерелом виступає файл MS Excel або текстовий файл.



Перш за все, для проведення імпорту даних необхідно підготувати файл. Дані мають бути структуровані та підписані. Слід пам'ятати, що при імпорті файлу MS Excel він має бути записаний у версії *Книга Excel 5.0/95*<sup>1</sup>. Приклад оформлення даних показаний на рис. 7, на якому показані дані ВВП України (*gdp*), рівня інвестицій (*inv*), державні витрати (*g*), грошова маса M2 (*m2*), облікова ставка НБУ (*r*).



	A	B	C	D	E	F	G	H
	Time	GDP	Inv	G	M2	R		
1								
2	1998/Q1	20871	1744	6720	12835	35		
3	1998/Q2	23367	2675	7384	13257	41		
4	1998/Q3	28908	2877	8210	14142	51		
5	1998/Q4	29447	6662	8952	15432	82		
6	1999/Q1	24980	1861	6245	15631	60		
7	1999/Q2	29196	3006	7737	18258	60		
8	1999/Q3	37633	4023	9672	20019	45		
9	1999/Q4	35317	8662	9677	21714	45		
10	2000/Q1	32309	2659	7980	23275	32		
11	2000/Q2	37889	4018	10609	26359	29		
12	2000/Q3	51238	5073	13271	28076	27		
13	2000/Q4	48634	11879	13763	31544	27		
14	2001/Q1	39201	3945	10506	32531	25		
15	2001/Q2	46481	6062	12782	36552	19		

Рис. 7. Приклад оформлення даних у MS Excel для імпортування в EViews

Наступним кроком є використання у EViews команди меню **File→Import→Read Text-Lotus-Excel** та обрання записаного файлу. У новому вікні вказується порядок розташування даних (по рядках чи стовпчиках), верхню ліву чарунку, де починаються безпосередньо дані (як правило, B2), при необхідності – назву листа книги MS Excel та вибірку, до якої слід імпортувати дані. Найважливішим кроком є перелік назв змінних, що мають бути імпортовані. В нашому прикладі мають бути імпортовані всі 5 змінних (рис. 8).

<sup>1</sup> Для версії EViews 5.0 і вище таке обмеження не діє: файл може бути записаний у будь-якій версії MS Excel.



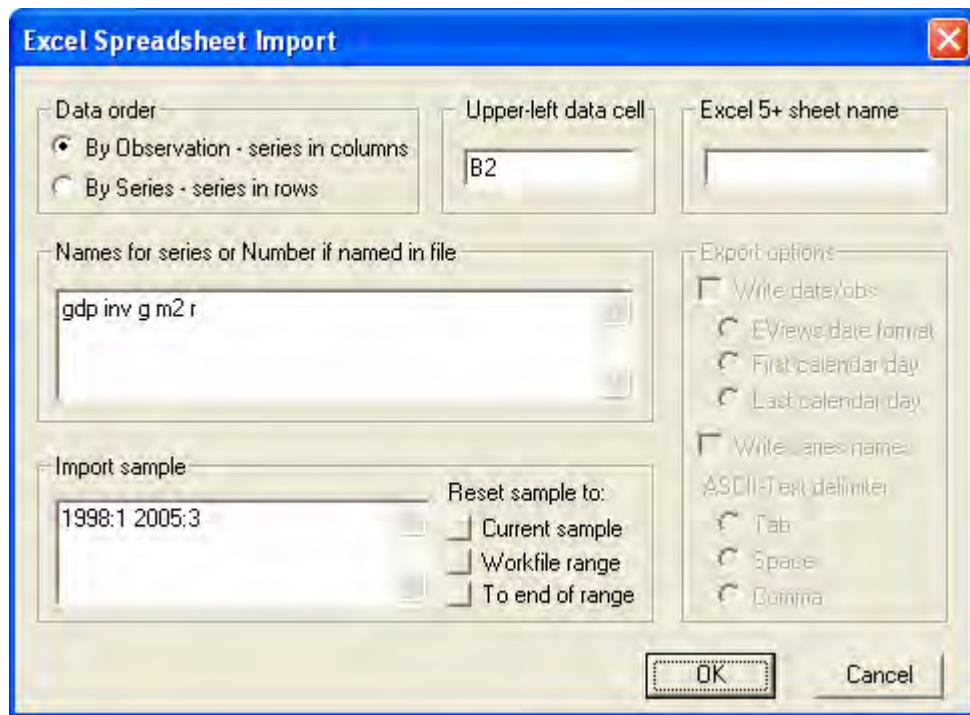


Рис. 8. Імпорт даних до EViews.

Перед натисканням кнопки *Ok* слід обов'язково впевнитися, що відповідний файл MS Excel закритий, інакше системою буде видана помилка й імпорт не буде здійснений. Якщо імпорт відбувся успішно, в робочій книзі з'являться 5 нових змінних. Аналогічно відбувається імпорт з інших джерел.

Третя можливість створення змінних передбачає генерацію змінної за допомогою формули. Для цього використовується команда меню **Quick→Generate Series...** або кнопка **Genr**. У вікні слід вказати вибірку, для якої створюється нова змінна, а також формулу, за якою будуть генеруватися елементи. На рис. 9 показано створення змінної, що обраховує частку інвестицій у загальному рівні ВВП країни. В результаті буде створена змінна *myratio*, кожне значення якої є часткою інвестицій у ВВП України у кожному кварталі протягом 1998-2005 років.

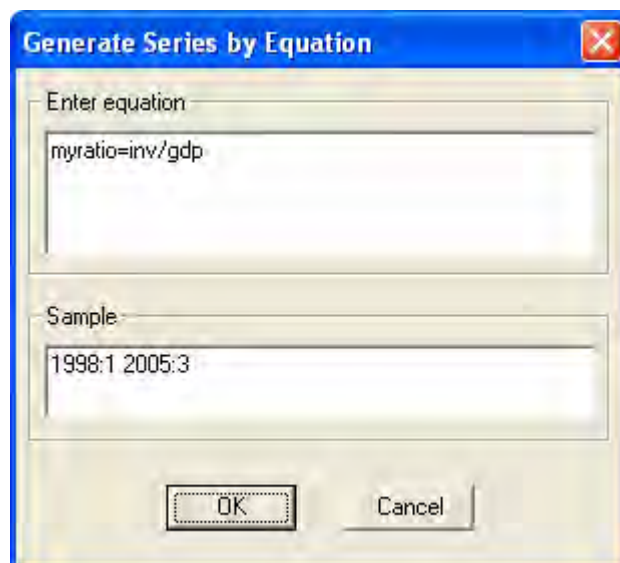


Рис. 9. Створення нової змінної за допомогою формули

В якості виразів у формулах можуть використовуватися:

- арифметичні оператори (+, -, \*, /, ^);
- оператори порівняння (<, >, <=, >=, =);
- логічні оператори (*and*, *or*);
- константи, існуючі змінні;
- функції системи.

Виділення декількох об'єктів здійснюється натисканням лівої кнопки мишки при натиснутій клавіші *Ctrl*. Знищення непотрібних об'єктів здійснюється за допомогою кнопки **Delete** або команди меню **Objects→Delete selected**.

## Бази даних

База даних у EViews (*Database*) являє собою файл, у якому містяться різні об'єкти. При роботі з базою даних до пам'яті комп'ютера завантажуються лише необхідні об'єкти, в той час як при роботі з робочою книгою завантажуються всі елементи. Така можливість дозволяє одночасно обробляти величезну кількість змінних, моделей, систем. Другою перевагою баз даних над звичайними робочими книгами є можливість зберігання змінних з різною періодичністю.

На диску база даних займає декілька файлів. Основні об'єкти містяться у файлі з розширенням *.edb*, файли індексації мають розширення *.el0*, *.el1a* тощо. При переносі бази даних з однієї папки до іншої чи іншого комп'ютера слід копіювати всі файли.

Оскільки вся база даних знаходиться на диску, то всі зміни відразу вносяться то відповідних файлів без можливості повернутись до останньої версії файлу.

Створення бази даних здійснюється аналогічно до створення робочої книги (**File→New→Database...**). Оскільки база даних має бути відразу створена на диску (рис. 10), то необхідно вказати її точне розташування (поле **DB File name/path**).

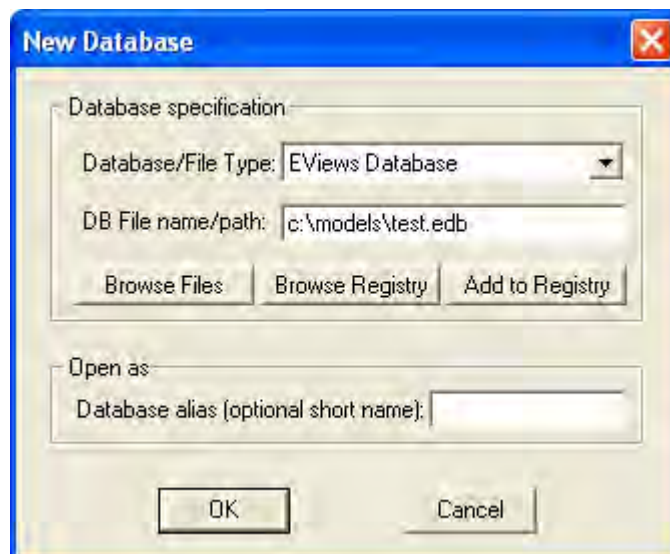


Рис. 10. Створення нової бази даних

Для перенесення до створеної бази даних об'єктів робочої книги використовується кнопка **Store** або команда контекстного меню об'єкта **Store to DB....** Обернена операція копіювання об'єкта бази даних до робочої книги здійснюється за допомогою кнопки **Fetch**, при цьому вказуються назви об'єктів, які необхідно імпортувати. Аналогічну операцію можна здійснити з вікна бази даних (кнопка **Export**).

Для управління розташуванням об'єктів бази даних можна застосовувати різноманітні фільтри (кнопки **All**, **EasyQuery**, **Query**). Меню **Procs** містить команди для копіювання бази даних, її перевірки на цілісність, відновлення пошкоджених об'єктів.

### ***Лабораторна робота***

**Задача 1.1.** Створити базу даних у основних макроекономічних показників України за 1996-2006 роки:

- ВВП;
- експорт та імпорт товарів та послуг;
- грошова маса M0, M1, M2, M3;
- рівень інфляції;
- курс гривні відносно основних світових валют;
- обсяг вкладів населення в банках.

Для подальшої роботи вибрати 5 часових рядів, що не повторюються з іншими студентами. Кожен з рядів має містити не менше 30 спостережень. Для визначення потрібної інформації скористатися сайтами:

- [www.me.gov.ua](http://www.me.gov.ua),
- [www.bank.gov.ua](http://www.bank.gov.ua).

**Задача 1.2.** Імпортувати базу даних до EViews.

**Задача 1.3.** Створити на основі даних зад. 1.2 файл бази даних у EViews. Використати її для розв'язання задач наступних розділів.

**Задача 1.4.** Використовуючи дані задачі 1.1. визначити

- рівень монетизації економіки (відношення грошової маси до ВВП);
- частку інвестицій у валовому продукті;
- динаміку приросту ВВП;
- динаміку приросту вкладів населення;
- частку експорту товарів та послуг у ВВП.

## РОЗДІЛ 2. ПРОГНОЗУВАННЯ ЗА ДОПОМОГОЮ ЧАСОВИХ РЯДІВ

### Теоретична частина

#### Основні визначення

В сучасній статистичній теорії існує багато різноманітних методів прогнозування економічної інформації. Значна їх частина відноситься до прогнозування часових рядів. Особливістю прогнозування часових рядів є те, що аналізуються лише дані спостережень, як правило, без додаткової інформації.

Нехай  $y_1, y_2, \dots, y_T$  - значення спостережень за економічним процесом протягом  $T$  періодів. Ця послідовність є числовими значеннями, кожне з яких має відповідний індекс, який залежить від номера періоду, в який він спостерігався. Така послідовність, записана у порядку зростання індексу, називається **часовим рядом**. Будемо позначати часовий ряд з  $T$  елементами  $\{Y_T\}$ .

Процес може бути **неперервним**, наприклад, зміна температури повітря в даній місцевості, або **дискретним**, наприклад, випуск автомобілів автозаводом. Внаслідок того, що абсолютна більшість всієї економічної інформації є дискретною, то будемо розглядати лише дискретні часові ряди.

За своєю природою, рівні часових рядів можуть бути моментними й агрегованими. Часовий ряд **моментний**, якщо значення процесу формується в даний момент (наприклад – температура повітря в даному районі в даний момент). Часовий ряд **агрегований**, якщо значення рівня утворюються шляхом агрегування (підсумовування) за певний період. Наприклад, кількість опадів за травень у даній місцевості, випуск продукції підприємством за рік тощо.

#### Порядок аналізу часових рядів

Метою прикладного статистичного аналізу часових рядів є побудова математичної моделі ряду, за допомогою якої можна пояснити поведінку ряду і здійснити прогноз на майбутні періоди.

**Побудова і вивчення графіка.** Аналіз часового ряду починається з побудови і вивчення його графіка. Процес виділення тренду та інших компонентів ряду, що призводять до порушення стаціонарності, може проходити в декілька етапів. На кожному з них розглядається ряд залишків, отриманий у результаті вирахування з вихідного ряду підібраної моделі тренду, або результат різницьових і інших перетворень ряду. Крім графіків, ознаками нестаціонарності часового ряду можуть служити певний вигляд автокореляційної та часткової автокореляційної функцій, числові значення

найпростіших числових характеристик (середнє, медіана, дисперсія, стандартне відхилення, коефіцієнти асиметрії й ексцесу).

Графічний аналіз ряду, як правило, задає напрямок його подальшого аналізу.

**Вибір моделі для часового ряду.** Мета цього етапу – опис і урахування надалі аналізу кореляційної структури аналізованого процесу. Модель може вважатися підбраною, якщо залишковий компонент ряду є процесом, як правило, "білого шуму". Після підбору залишки аналізуються для перевірки адекватності моделі та побудови надійних інтервалів.

**Прогнозування або інтерполяція.** Останнім етапом аналізу часового ряду може бути прогнозування його майбутніх (екстраполяція) або відновлення пропущених (інтерполяція) значень і визначення точності цього прогнозу на базі підбраної моделі. Добре підібрати математичну модель вдається не для всякого часового ряду. Нерідко буває і так, що для опису підходять відразу декілька моделей. Неоднозначність вибору моделі може спостерігатися як на етапі виділення детермінованого компонента ряду, так і при виборі структури ряду залишків. Тому досить часто розробляють декілька прогнозів, зроблених за допомогою різних моделей.

### Критерії точності прогнозів

Найбільш простою мірою якості прогнозів за умови, що є дані про їхню реалізацію, може стати відносне число випадків, коли фактична реалізація попадала у довірчий інтервал прогнозу, до загального числа прогнозів, тобто

$$\eta = \frac{m}{m + n},$$

де  $m$  – кількість прогнозів, підтверджених фактичними даними;

$n$  – кількість прогнозів, не підтверджених фактичними даними. Коли всі прогнози підтверджуються,  $n = 0$  і  $\eta = 1$ , якщо ж усі прогнози не підтвердилися, то  $m$  і  $\eta$  рівні 0.

Одним із дослідників проблем економічного прогнозування, Г. Тейлом, запропонований за міру якості прогнозу коефіцієнт розбіжності (або коефіцієнт невідповідності), чисельником якого є середньоквадратична помилка прогнозу, а знаменник дорівнює квадратному кореню із середнього квадрата реалізації, тобто:

$$v = \frac{\sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}}{\sqrt{\frac{\sum_{t=1}^n y_t^2}{n}}} = \frac{\sqrt{\sum_{t=1}^n (\hat{y}_t - y_t)^2}}{\sqrt{\sum_{t=1}^n y_t^2}}.$$

Коефіцієнт  $v = 0$ , коли всі  $\hat{y}_t = y_t$  (випадок ідеального прогнозування);  $v = 1$ , коли процес прогнозування призводить до середньоквадратичної помилки "наївної" екстраполяції незмінності приростів; нарешті,  $v > 1$ , коли

прогноз дає гірші результати, ніж припущення про незмінність досліджуваного явища.

Коефіцієнт невідповідності може бути використаний при зіставленні якості прогнозів, одержаних на основі різноманітних методів і моделей, у чому полягає його безсумнівна привабливість.

Іноді коефіцієнт Тейла розраховують через середньоквадратичне значення похибки прогнозу приростів:

$$U = \sqrt{\frac{\frac{1}{n} \sum_{t=1}^n (\Delta \hat{y}_{T+t} - \Delta y_{T+t})^2}{\frac{1}{n} \sum_{t=1}^n (\Delta y_{T+t})^2}} = \sqrt{\frac{\sum_{t=1}^n (\Delta \hat{y}_{T+t} - \Delta y_{T+t})^2}{\sum_{t=1}^n (\Delta y_{T+t})^2}}.$$

У стандартній статистиці для міри точності прогнозів частіше використовуються такі характеристики:

$$MSE = \frac{1}{n} \sum_t (y_t - \hat{y}_t)^2 - \text{середньоквадратична похибка прогнозу за } n \text{ кроків.}$$

$RMSE = \sqrt{\frac{1}{n} \sum_t (y_t - \hat{y}_t)^2}$  – корінь із середньоквадратичної похибки прогнозу за  $n$  кроків.

$$MAD = \frac{1}{n} \sum_t |y_t - \hat{y}_t| - \text{середня абсолютна похибка за } n \text{ кроків.}$$

$RMSPE = 100 \sqrt{\frac{1}{n} \sum_t \left( \frac{y_t - \hat{y}_t}{y_t} \right)^2}$  – корінь із середньоквадратичної похибки у відсотках від фактичних значень за  $n$  кроків.

$MAPE = \frac{100}{n} \sum_t \left| \frac{y_t - \hat{y}_t}{y_t} \right|$  – середня абсолютна похибка у відсотках за  $n$  кроків.

На практиці ці характеристики використовуються досить часто. Перші три критерії виражають похибку у одиницях виміру, тому їх величина залежить від специфіки часового ряду. Останні два критерії вимірюються у відносних одиницях, тому можна говорити про деякий загальний рівень адекватності моделі на основі їх порівняння. Чим меншою є величина критерію похибки, тим краще побудована модель для прогнозування.

### Аддитивна та мультиплікативна моделі часових рядів

Будь-який часовий ряд можна представити як суму детермінованого та випадкового компонентів:

$$y_t = d_t + r_t, \quad t = \overline{1, T}.$$

В свою чергу детермінований компонент складається з трьох частин: трендового, сезонного, циклічного компонентів.

$$d_t = tr_t + s_t + c_t, \quad t = \overline{1, T}.$$

Таким чином, будь-який часовий ряд можна розглядати як суму:

$$y_t = tr_t + s_t + c_t + r_t, \quad t = \overline{1, T}.$$

Такий вигляд часового ряду отримав назву **адитивної моделі**. Якщо ж замість реальних значень компонентів використовувати їх логарифми, то отримаємо **мультиплікативну модель**:

$$\ln y_t = \ln tr_t + \ln s_t + \ln c_t + r_t, \quad t = \overline{1, T}, \text{ або } y_t = tr_t \cdot s_t \cdot c_t + r_t, \quad t = \overline{1, T}.$$

**Детермінований компонент** змінюється за певними правилами, які можуть бути визначені за допомогою досліджень і відповідного аналізу часового ряду. Як правило, одним з основних параметрів, від яких залежить детермінований компонент, є час.

Аналіз часового ряду починається з виділення **трендового компонента**. Його присутність неважко помітити, проаналізувавши графік часового ряду. Як правило, для економічних даних дуже типовим є повільне зростання чи падіння протягом тривалого періоду часу. Наявність **тренду** в економічних часових рядах можна пояснити демографічними змінами, технологічними змінами, змінами в структурі виробництва, попиту, тощо. Дія таких факторів є постійною, тому дослідники мають змогу описувати такі зміни за допомогою кривих, які можна задати в аналітичному вигляді.

**Сезонний компонент** показує коливання навколо трендового компонента. Його наявність пояснюється сезонним характером виробництва, споживання. Наприклад, у четвертому кварталі кожного року перед Новим роком значно зростає споживання товарів. Головна ідея виділення сезонних коливань полягає у порівнянні даних за відповідні періоди, а не за минулі періоди, тобто, наприклад, дані за грудень одного року треба порівнювати з даними грудня минулих років, а не з листопадом.

**Циклічний компонент** займає проміжне місце між трендом та сезонним компонентом. Тренд – це гладка зміна, яка проявляється на великому проміжку часу. Сезонний компонент – це періодична функція, що залежить від часу, причому його період значно менший за кількість спостережень. Циклічний компонент розглядається, в основному, як гладка зміна, залежна від часу, але яка не включається ні до тренду, ні до сезонного компонента.

**Випадковий компонент** є те, що залишилось від часового ряду після виключення тренду, циклічного та сезонного компонентів. Частина таких ефектів може бути віднесена до непередбачених природних катаклізмів (землетруси, пожежі, тощо), частина – до випадкових дій людей. За наявності випадкового компонента неможливо прогнозувати значення часового ряду без помилки.

### Числові характеристики часових рядів

Для будь-якого часового ряду можна знайти деякі числові характеристики. Для аналізу часових рядів найважливішими є математичне сподівання, дисперсія, коваріація, кореляція.

Математичне сподівання часового ряду  $\{Y_T\}$  є

$\mu_t = Ey_t = \int_{-\infty}^{\infty} x dF_t(x)$ , де  $F_t(x) = P\{y_t < x\}$  – функція розподілу  $y_t$ ,  $t = -\infty, \dots, \infty$ .

Дисперсія часового ряду  $\{Y_t\}$  визначається за формулою:

$$\text{var}(y_t) = E(y_t - Ey_t)^2.$$

Нарешті, автоковаріація часового ряду  $\{Y_t\}$  дорівнює:

$$\text{cov}(y_t, y_{t-j}) = E((y_t - Ey_t)(y_{t-j} - Ey_{t-j})), \quad j = 1, 2, 3, \dots$$

Для отримання практичних оцінок для часових рядів користуються формулами:

$$\text{математичне сподівання} - \hat{\mu} = \frac{1}{T} \sum_{t=1}^T y_t,$$

$$\text{дисперсія} - \hat{\gamma}_0 = \frac{1}{T-1} \sum_{t=1}^T (y_t - \hat{\mu})^2,$$

$$\text{автоковаріація } j\text{-го порядку} - \hat{\gamma}_j = \frac{1}{T-j} \sum_{t=j+1}^T (y_t - \hat{\mu})(y_{t-j} - \hat{\mu}).$$

Часовий ряд є **стаціонарним**, якщо

- 1) математичне сподівання  $Ey_t = \mu < \infty$  для всіх  $t$ ,
- 2) дисперсія  $\text{var}(y_t) = \gamma_0 < \infty$  для всіх  $t$ ,
- 3) автоковаріація  $j$ -го порядку  $\text{cov}(y_t, y_{t-j}) = \gamma_j < \infty$  для всіх  $t$ ,  
 $j = 1, 2, 3, \dots$

Звичайно, жоден з рядів, що представляє реальну економічну інформацію, не може бути ідеально стаціонарним. Але якщо для деякого часового ряду з деяким наближенням виконуються умови стаціонарності, то для його аналізу можна використати широкий спектр методів аналізу та прогнозування стаціонарних часових рядів.

Крім вищенаведених характеристик при аналіз часових рядів застосовується автокореляція та автокореляційна функція. Автокореляція  $j$ -го порядку визначається за допомогою рівності:

$$\rho_j = \frac{\text{cov}(y_t, y_{t-j})}{\text{var}(y_t)} = \frac{\gamma_j}{\gamma_0}.$$

Цей коефіцієнт визначає ступінь залежності між спостереженнями, які знаходяться на відстані  $j$  періодів.

Якщо побудувати залежність  $\rho_j$  на графіку, то отримаємо **корелограму**. Вона представляє деяку криву, що показує, як змінюється взаємовплив між спостереженнями в залежності від часу.



## Метод поворотних значень

Під **випадковим часовим рядом** будемо розуміти такий ряд, у якому значення є випадковими незалежними величинами, що мають один закон розподілу. Якщо вдається показати, що часовий ряд є випадковим, то подальше його дослідження можна припинити, вирахувавши, при необхідності, середнє значення та дисперсію ряду або інші статистичні характеристики.

Одним з методів перевірки часового ряду на випадковість є **метод поворотних значень**.

Для часового ряду  $\{Y_T\}$  значення  $y_i$  є поворотним, якщо  $y_{i-1} < y_i > y_{i+1}$  або  $y_{i-1} > y_i < y_{i+1}$ . У першому випадку значення  $y_i$  є «піком», у другому — «впадиною». Зауважимо, що перше й останнє значення не можуть бути поворотними.

Метод поворотних значень зводиться до підрахунку кількості впадин і піків у ряду  $\{Y_T\}$  і порівнянні цієї кількості з теоретичним значенням, яке дорівнює математичному сподіванню кількості поворотних точок у «чисто випадковому» ряду, що складається з  $T$  спостережень.

Нехай

$$x_i = \begin{cases} 1, & \text{якщо } y_{i+1} \text{ поворотне значення;} \\ 0, & \text{в інших випадках.} \end{cases}$$

Математичне сподівання кількості поворотних точок у випадковому ряді  $\{Y_T\}$ , у силу незалежності  $x_i$ , буде дорівнювати:

$$\mu(T) = E(k) = \sum_{i=1}^{T-2} E(x_i) = \frac{2}{3}(T-2),$$

середнє квадратичне відхилення:

$$\sigma(k) = \sqrt{\frac{16T-29}{90}}.$$

Для перевірки гіпотези про випадковість ряду використовується значення:

$$U_{pr} = \frac{|k_{pr} - \mu(T)|}{\sigma(T)},$$

яке порівнюється з теоретичним  $U_{teor}$ . Якщо  $U_{pr} < U_{teor}$ , то немає підстави відхилити нульову гіпотезу, тобто ми повинні прийняти твердження, що даний ряд випадковий. У протилежному випадку, коли  $U_{pr} > U_{teor}$ , треба прийняти твердження про невідповідність ряду.

Для найбільш поширених значень  $\alpha$  відповідні величини  $U_{kr}$  наведені у таблиці.

$\alpha$	$H^1 : k_{pr} \neq \mu(T)$	$H^1 : k_{pr} < \mu(T)$	$H^1 : k_{pr} > \mu(T)$
0,01	2,58	2,33	2,33
0,05	1,96	1,645	1,645
0,1	1,645	1,283	1,283

## Практична частина

### Робота зі змінними

З кожною змінною EViews, а іноді і з групою змінних, можна проводити різноманітні операції. Зокрема, можна вивести статистичну інформацію по змінній за допомогою меню **View→Descriptive Statistics→Histogram and Stats** (рис.11). В результаті цього програма виводить інформацію про найбільше та найменше значення змінної, вибіркове середнє, медіану ряду, середньоквадратичне відхилення, ступінь асиметричності ряду, виводить статистику Жарку-Бера, що перевіряє гіпотезу про нормальний розподіл ряду, імовірність прийняття гіпотези (наприклад, на рис. 11 гіпотеза про нормальний розподіл приймається).

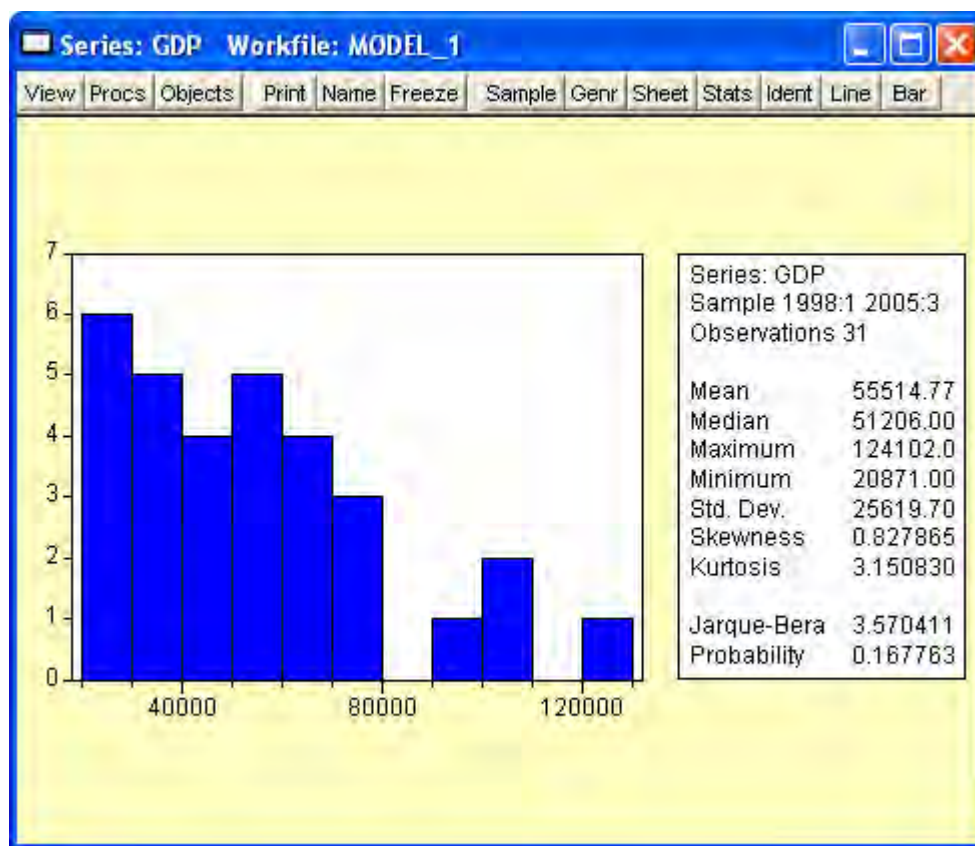


Рис. 11. Вивід основних статистичних характеристик змінної

Аналогічну інформацію, але у табличному виді видає команда меню **View→Descriptive Statistics→Stats Table**. Нарешті, команда **View→Descriptive Statistics→Stats by Classification** дозволяє виводити статистику для довільних підвибірок заданого ряду, враховуючи різноманітні обмеження та умови.

За допомогою меню **View→Tests for Descriptive Stats→Simple Hypothesis Tests** здійснюється перевірка гіпотез про параметри розподілу ряду (середнє, дисперсію, медіану). За допомогою тесту можна визначити, наприклад, чи дорівнює середнє значення ряду 60 000 (рис. 12).

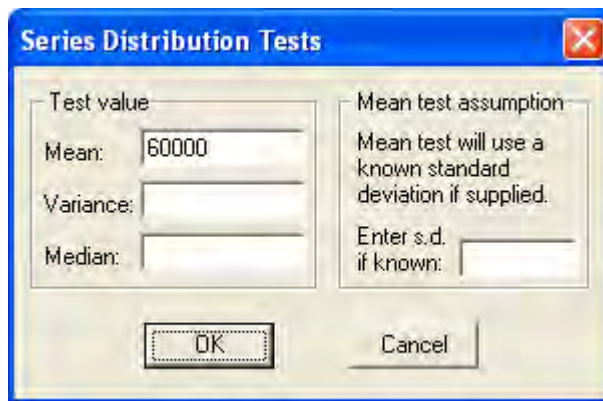


Рис. 12. Проведення тестування середнього значення ряду

Результати обчислень (рис. 13) свідчать, що гіпотеза про те, що вибіркове середнє дорівнює 60000, має бути прийнята, оскільки імовірність прийняття гіпотези 33,75% значно більше необхідного рівня (як правило, 5%). Таким чином, можна вважати, що середнє значення ряду *gdp* дорівнює 60 000.

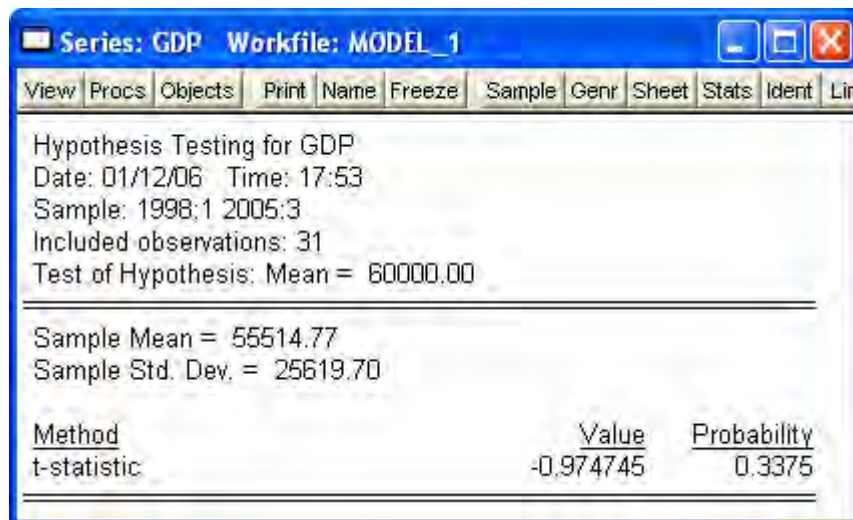


Рис. 13. Результати тестування

Команда **View→Tests for Descriptive Stats→ Equality Tests by Classification...** виконує аналогічні дії, але з підвибірками основного ряду.

Меню **View→Distribution** дозволяє побудувати різноманітні графіки, що характеризують емпіричний розподіл часового ряду. Підменю **CDF-Survivor-Quantile...** може вивести емпіричну функцію розподілу, емпіричні квантілі розподілу. Підменю **Quantile-Quantile Graphs...** дозволяє порівняти 2 розподіли. Найчастіше ця опція використовується для графічного порівняння заданого часового ряду з теоретичним розподілом (нормальним, рівномірним тощо). Підменю **Kernel Density Graphs...** дозволяє побудувати графік функції щільності часового ряду. Нарешті, підменю **Empirical Distribution Tests** дозволяє перевірити гіпотезу про відповідність часового ряду заданому розподілу. Тестування ґрунтується на порівнянні емпіричної та теоретичної функцій розподілу. Застосуємо тест для перевірки нормального розподілу змінної *gdp*. На рис. 14 представлені результати перевірки. За чотирма

методами гіпотеза про нормальний розподіл приймається, що можна побачити по значеннях стовпчика *Probability*: всі значення перевищують стандартний рівень 0,05. Також EViews оцінив параметри розподілу: середнє та середньоквадратичне відхилення (значення *mu* та *sigma* відповідно).

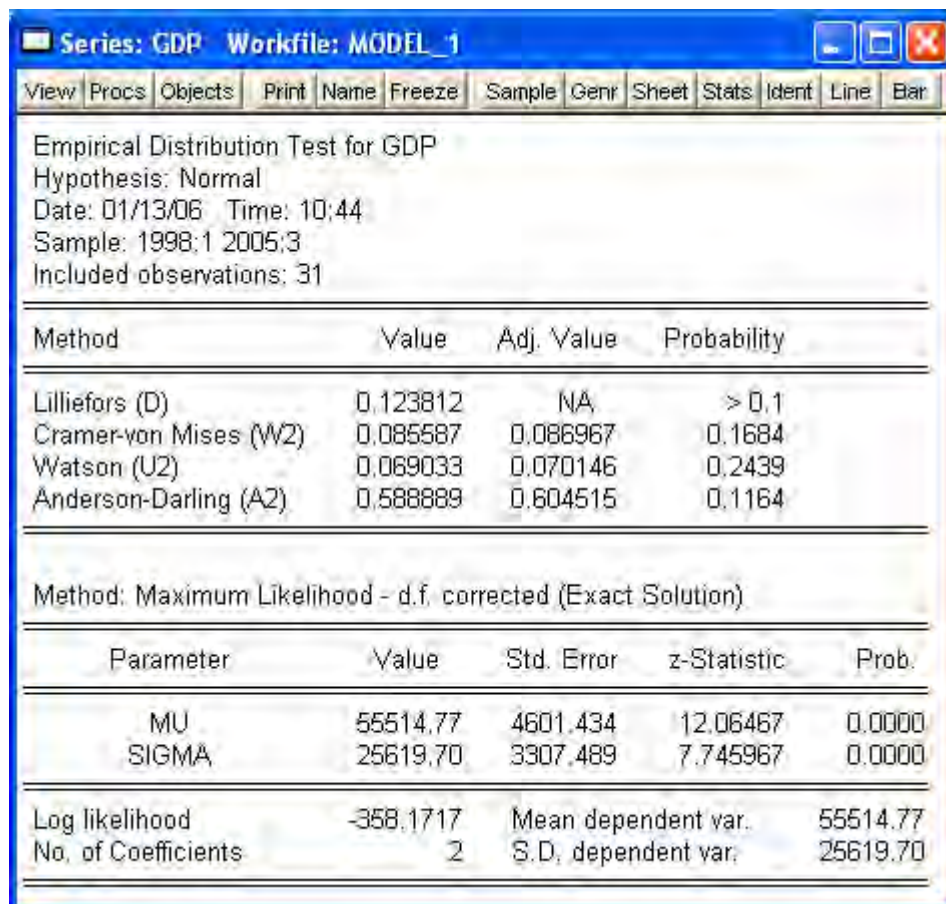


Рис. 14. Перевірка гіпотези про нормальний розподіл змінної GDP.

Меню **View→One-Way Tabulation...** дозволяє згрупувати дані ряду по групах у порядку зростання, показуючи інформацію про частоти кожної групи та накопичені частоти.

Важливим для аналізу ряду є побудова корелограми за допомогою меню **View→Correlogram...**. При цьому вказується, для яких саме значень змінної слід робити розрахунки: для самої змінної, для перших різниць, для других різниць. Кількість лагів визначає розмір самої корелограми. Наприклад, для змінної ВВП графік кореляційної та часткової кореляційної функцій матиме вигляд, представлений на рис. 15.

Штрихові лінії на графіках показують надійний інтервал, в якому значення статично приймається рівним 0. Числові характеристики функції кореляції та часткової кореляційної функції знаходяться у стовпчиках *AC* та *PAC* відповідно. Значення *Q-Stat* місять значення статистики Лjung-Бокса (див. розділ 5) для перевірки гіпотези про те, що автокореляція відсутня до *k*-го лагу включно. В нашому випадку така гіпотеза відхиляється для всіх лагів від 1 до 16 (значення *Prob*<0.05).



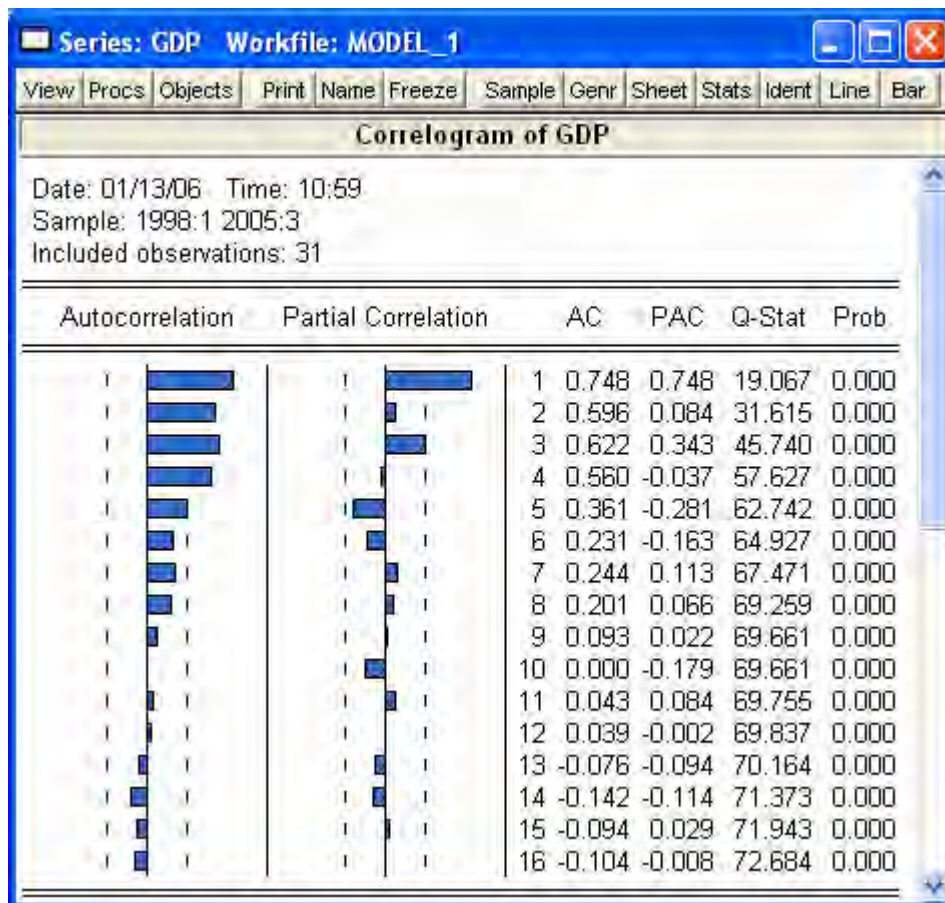


Рис. 15. Графік кореляційної та часткової кореляційної функції для змінної GDP.

Тести на стаціонарність (**View→Unit Root Test...**) та незалежність (**View→BDS Independence Test...**) будуть розглянуті пізніше (див. розділ 5).

Меню **View→Label** дозволяє вводити допоміжну інформації щодо змінної. При наявності великої кількості об'єктів зручно відображати таку інформацію за допомогою перемикача **Label+/-** на панелі робочої книги або за допомогою меню **View→Display Comments (Label+-)**.

### Групові операції

Обрання декількох змінних утворює групу. Її можна записати як один об'єкт (позначається символом **G**), що позбавить необхідності вибирати кожен раз одні й ті ж дані. Об'єкти-групи позначаються символом. Над елементами групи можна здійснювати операції паралельного редагування змінних, періодичного (поквартальне, помісячне тощо) порівняння даних (**View→Dated Data Table**), побудови різноманітних типів діаграм, проведення тестів тощо.

Існує можливість одночасного графічного зображення декількох змінних. При застосуванні команди **View→Show...** виводиться табличне зображення. Для переходу у графічний режим слід вибрати **View→Graph→Line**. Натиснення правою кнопкою миші на графіку дозволяє змінити опції графіка, налаштувавши його на свій смак.

## **Лабораторна робота**

**Задача 2.1.** Провести графічний аналіз рядів даних. Визначити наявність трендового, сезонного компонентів.

**Задача 2.2.** Для обраних рядів даних визначити основні числові характеристики:

- вибіркове середнє;
- вибіркиму дисперсію;
- коваріацію перших 10 порядків.

Побудувати гістограми часових рядів.

**Задача 2.3.** Побудувати корелограму та часткову кореляційну функцію для рівнів часових рядів, їх перших та других різниць.

**Задача 2.4.** Перевірити Ваші часові ряди на випадковість за допомогою методу поворотних значень.

**Задача 2.5.** Розбити всі ряди на дві однакові за розмірами вибірки. Перевірити гіпотезу про

- рівність математичних сподівань;
- дисперсій

у вибірках.

**Задача 2.6.** Створити на основі змінних бази даних нові, що являють собою перші різниці відповідних часових рядів. На основі отриманих значень перевірити гіпотези:

- про випадковість за допомогою методу поворотних значень;
- про нормальний розподіл за допомогою декількох методів.

**Задача 2.7.** Обчислити ряди других різниць початкових даних. Перевірити гіпотези:

- про випадковість за допомогою методу поворотних значень;
- про нормальний розподіл за допомогою декількох методів;
- про рівність вибіркового середнього 0.

**Задача 2.8.** Для „наївної” моделі (прогноз дорівнює останньому значенню часового ряду) підрахувати помилки прогнозування за критеріями:

- MSE;
- RMSE;
- MAD;
- RMSPE;
- MAPE;
- коефіцієнтом Тейла.

## РОЗДІЛ 3. МЕТОДИ ЗГЛАДЖУВАННЯ ЧАСОВИХ РЯДІВ

### Теоретична частина

Методи згладжування використовуються для зменшення впливу випадкового компонента (випадкових коливань) у часових рядах. Вони дають можливість отримувати більш “чисті” значення, які складаються лише з детермінованих компонентів. Деякі з методів направлені на виділення лише деяких компонентів, наприклад, тренду.

#### Метод усереднення (moving average)

Цей метод є одним з найпростіших, який дозволяє виділити тренд. Для застосування цього методу дослідник повинен мати доволі довгий ряд спостережень. Формально метод описується виразом:

$$S_t = \frac{1}{k} \sum_{j=-k_1}^{k_2} y_{t+j}, \quad k = k_1 + k_2 + 1.$$

Для квартальних даних часового ряду при  $k_1 = 3$ ,  $k_2 = 0$  формула набуває вигляду

$$S_t = \frac{1}{4} \sum_{j=-3}^0 y_{t+j} = \frac{1}{4} (y_{t-3} + y_{t-2} + y_{t-1} + y_t).$$

З формули видно, що нова кількість спостережень становить  $T - k$ . За допомогою цього методу можна не тільки більш чітко спостерігати трендовий компонент, але й сезонні та випадкові коливання. Для цього треба використовувати мультиплікативну модель часового ряду:

$$S_t = tr_t \cdot c_t, \quad t = \overline{1, T - k}.$$

Тоді

$$\frac{y_t}{S_t} = \frac{tr_t \cdot c_t \cdot s_t \cdot r_t}{tr_t \cdot c_t} = s_t \cdot r_t, \quad t = \overline{1, T - k}.$$

Єдиною складністю є визначення чисел  $k_1$  та  $k_2$ . Як правило, їх сума дорівнює повному циклу сезонності, тобто, наприклад, для квартальних даних – року.

Більш гладкий тренд дозволяє виділити метод подвійного усереднення, яке двічі використовує усереднення часового ряду. При цьому кількість спостережень зменшується на два повних цикли сезонності, тому для використання методу необхідно мати часовий ряд, який складається щонайменше з 3-х повних циклів сезонності.

#### Звичайне експоненціальне згладжування

Цей метод значно переважає усі попередні моделі. Найкраще цей метод зарекомендував себе, коли дані мають дуже гладкий, або навіть горизонтальний тренд. Нова послідовність будується за правилом:

$$S_1 = y_1,$$

$$S_t = \alpha y_t + (1 - \alpha) S_{t-1}, \quad t = \overline{2, T} \quad 0 < \alpha < 1.$$

Єдина вага  $\alpha$  може обиратися кількома шляхами. По-перше, якщо обирається значення близьке до 1, то будуть більш важливими при прогнозуванні останні дані часового ряду, при виборі  $\alpha$  близьким до 0, більш впливовими будуть минулі значення. По-друге, можна покласти  $\alpha = \frac{2}{T+1}$ . По-третє – вибір  $\alpha$ , при якому мінімізується один з критеріїв точності прогнозів на  $n$  періодів. При цьому розрахунки повинні проводитися лише по перших  $T - n$  значеннях часового ряду, а отримані прогнози повинні бути порівняні з реальними даними. При використанні цієї методики відкидається обмеження  $0 < \alpha < 1$ .

Прогноз значень часового ряду дорівнює останньому члену послідовності  $S_t$ :

$$\hat{y}_{T+p} = S_T, \quad p = 1, 2, \dots$$

### Подвійне експоненціальне згладжування Брауна

Цей метод будується аналогічно попередньому, тільки процес згладжування робиться двічі:

$$S_1 = y_1,$$

$$S'_t = \alpha y_t + (1 - \alpha) S'_{t-1},$$

$$S''_t = \alpha S'_t + (1 - \alpha) S''_{t-1}, \quad t = \overline{2, T} \quad 0 < \alpha < 1.$$

Метод використовується, коли дані часового ряду нестационарні. Прогноз будується як останнє значення другої послідовності:

$$\hat{y}_{T+p} = S''_T, \quad p = 1, 2, \dots$$

### Потрійне експоненціальне згладжування Брауна

Цей метод аналогічний двом попереднім, тільки згладжування проводиться тричі. Це дозволяє прогнозувати нестационарні часові ряди з великими перепадами мінімального та максимального значень. Нові послідовності будуються за правилом:

$$S'_t = \alpha y_t + (1 - \alpha) S'_{t-1},$$

$$S''_t = \alpha S'_t + (1 - \alpha) S''_{t-1},$$

$$S'''_t = \alpha S''_t + (1 - \alpha) S'''_{t-1}, \quad 0 < \alpha < 1.$$

Прогноз на наступні періоди має вигляд:

$$\hat{y}_{T+p} = S'''_T, \quad p = 1, 2, \dots$$

### Адаптивне згладжування

Цей метод дозволяє автоматично змінювати константу згладжування в процесі обчислення. Нова послідовність будується за правилом:



$$S_{t+1} = \alpha_t y_t + (1 - \alpha_t) S_t,$$

де  $\alpha_t$  змінюється з часом в залежності від похибки прогнозування:

$$\alpha_t = \left| \frac{E_t}{M_t} \right|,$$

а

$$E_t = \beta (y_t - \hat{y}_t) + (1 - \beta) E_{t-1},$$

$$M_t = \beta |y_t - \hat{y}_t| + (1 - \beta) M_{t-1}.$$

Параметр  $\beta$  знаходиться у межах  $(0; 1)$ . Для випадкових похибок коефіцієнт  $\alpha$  буде близьким до 0.5.

### Несезонна модель Холта-Вінтерса

Ця модель схожа на подвійне експоненціальне згладжування, але дозволяє виділяти трендовий компонент за допомогою другої послідовності:

$$S'_2 = y_2, \quad S''_2 = y_2 - y_1,$$

$$S'_t = \alpha y_t + (1 - \alpha) (S'_{t-1} + S''_{t-1}), \quad 0 < \alpha < 1,$$

$$S''_t = \beta (S'_t - S'_{t-1}) + (1 - \beta) S''_{t-1}, \quad 0 < \beta < 1.$$

Прогноз на наступні періоди:

$$\hat{y}_{T+p} = S'_T + p S''_T, \quad p = 1, 2, \dots$$

Слід відмітити, що прогнози, зроблені за цим методом, як правило, або сильно завищені, або занижені внаслідок того, що додається постійний трендовий компонент, який на практиці змінюється протягом року.

### Адитивна модель Холта-Вінтерса

Адитивна модель Холта-Вінтерса є розвиненням експоненціального згладжування. Аналізуються три статистично залежні ряди, які використовуються для побудови дійсного прогнозу: згладжені дані, трендовий компонент, індекс сезонності. Згладжена послідовність будується за правилом:

$$S_{t+p} = a_t + b_t p + c_{t+p},$$

де

$$a_t = \alpha (y_t - c_{t-s}) + (1 - \alpha) (a_{t-1} + b_{t-1}),$$

$$b_t = \beta (a_t - a_{t-1}) + (1 - \beta) b_{t-1},$$

$$c_t = \gamma (y_t - a_t) + (1 - \gamma) c_{t-s}, \quad 0 \leq \alpha, \beta, \gamma \leq 1.$$

Прогнози обраховуються за формулою:

$$\hat{y}_{T+p} = a_T + b_T p + c_{T+p-s}.$$

### Мультиплікативна модель Холта-Вінтерса

Мультиплікативна модель Холта-Вінтерса використовує індекс сезонності:

$$S_{t+p} = (a_t + b_t p) c_{t+p},$$

де

$$\begin{aligned}a_t &= \alpha \left( \frac{y_t}{c_{t-s}} \right) + (1 - \alpha)(a_{t-1} + b_{t-1}), \\b_t &= \beta (a_t - a_{t-1}) + (1 - \beta)b_{t-1}, \\c_t &= \gamma \left( \frac{y_t}{a_t} \right) + (1 - \gamma)c_{t-s}, \quad 0 \leq \alpha, \beta, \gamma \leq 1, \quad t = \overline{2s+1, T},\end{aligned}$$

де  $s$  - кількість циклів сезонності. Цей метод вимагає щонайменше 2 повних цикли сезонності по минулих спостереженнях для побудови розрахунків.

Прогноз на період  $T + p$  будується наступним чином:

$$\begin{aligned}\hat{y}_{T+p} &= (y_T + pb_T)c_{T-s+p}, \quad p = 1, 2, \dots, s, \\ \hat{y}_{T+p} &= (y_T + pb_T)c_{T-2s+p}, \quad p = s+1, s+2, \dots, 2s.\end{aligned}$$

### Фільтр Ходріка–Прескотта

Фільтр Ходріка-Прескотта (Hodrick-Prescott Filter) використовується для виділення трендового компонента. Нехай  $y_t = f(t) + \varepsilon_t$ , де  $y_t$  - значення змінної, що досліджується,  $t$  - час;  $\varepsilon_t$  - похибка;  $f(t)$  - деяка функція, яка може залежати від декількох параметрів, а також обов'язково залежить від часу.

Параметри у функції  $f$  підбираються таким чином, щоб мінімізувати вираз:

$$S = \sum_{t=1}^T (y_t - f(t))^2 + \lambda \sum_{t=2}^{T-1} ((f(t+1) - f(t)) - (f(t) - f(t-1)))^2 \rightarrow \min.$$

Практично було винайдено, що для річних даних прийнятним значенням вагового коефіцієнта є  $\lambda = 100$ , для квартальних -  $\lambda = 1600$ , для місячних -  $\lambda = 14400$ .

Отримані в результаті виконання фільтру залишки,  $\varepsilon$ , як правило, стаціонарним процесом. Це означає, що вони піддаються опису за допомогою деякого  $AR(p)$ -,  $MA(q)$ - чи  $ARMA(p, q)$ -процесу.

## Практична частина

### Методи згладжування економічної інформації

Для застосування методів згладжування слід відкрити необхідну змінну та застосувати команду **Procs**→**Exponential Smoothing**.... Відкрите вікно (рис. 16) містить 5 блоків.

У верхньому лівому блоці слід вибрати метод згладжування. EViews пропонує п'ять методів: звичайне експоненціальне згладжування (*Single*), подвійне експоненціальне згладжування (*Double*), несезонну модель Холта-Вінтерса (*Holt-Winters – No seasonal*), адитивну модель Холта-Вінтерса (*Holt-Winters – Additive*), мультиплікативну модель Холта-Вінтерса (*Holt-Winters – Multiplicative*). З правого боку показана кількість констант, що використовуються при оцінці кожного з методів.

Значення констант можна вказати вручну у нижньому лівому блоці вікна. Якщо значення констант не вказати (залишити знак *E*), то EViews автоматично обрахує їх значення за допомогою мінімізації критерію похибки *RMSE*.

Верхній правий блок містить назву нової послідовності, яка буде заповнена при обчисленнях. Нижче вказується вибірка, за якою слід проводити обчислення.

Нарешті, нижній правий блок містить число періодів в сезоні. Автоматично EViews пропонує значення 4 для квартальний даних, 12 – для місячних. Зауважимо, що це число використовується лише у адитивній та мультиплікативній моделях Холта-Вінтерса.

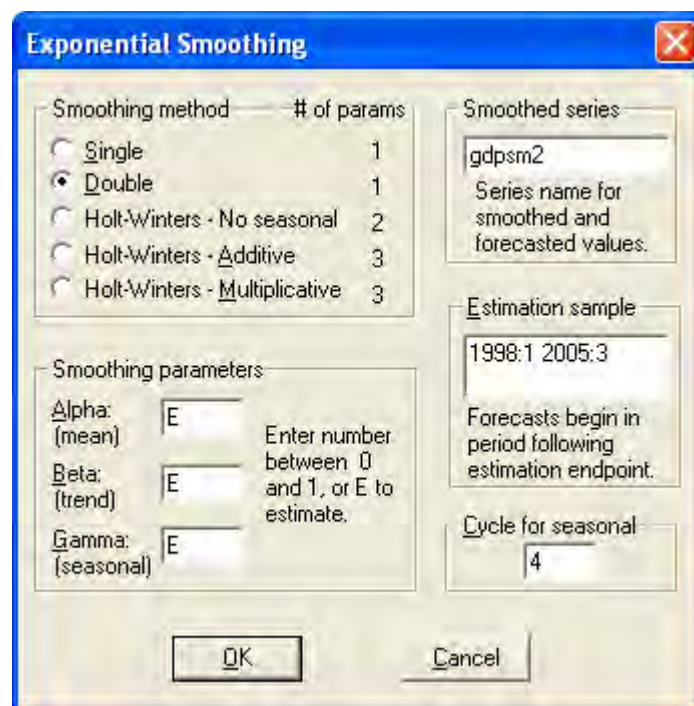


Рис. 16. Побудова методів згладжування

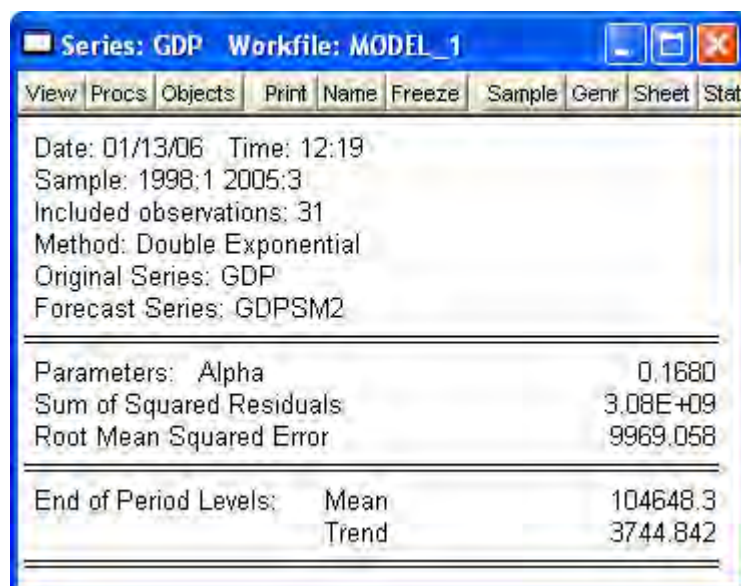


Рис. 17. Результати обчислень для подвійного експоненціального згладжування

Виконання команди створює в системі нову змінну *gdpsm2* та відкриває вікно результатів обчислень (рис. 17).

EViews обрахував значення константи  $Alpha=0.168$ , при якому критерій похибки RMSE приймає свого найменшого значення 9969.058.

Часто для аналізу вихідна та обчислена послідовності будуються на одному графіку. Для цього слід виділити при нажатій клавіші *Ctrl* ряди *gdp* та *gdpsm2*, і застосувати команди **Show** та **View→Graph→Line** (рис.18).

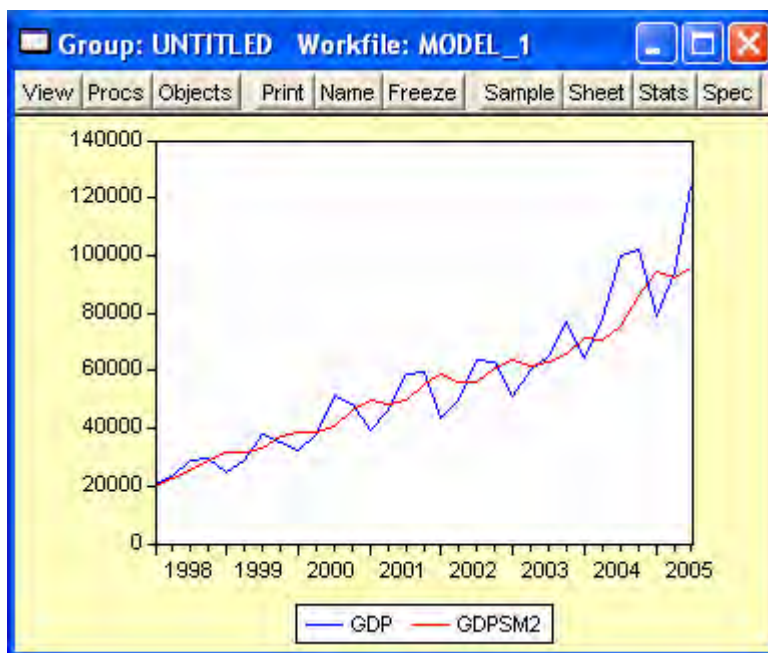


Рис. 18. Динаміка рівня ВВП та згладженої послідовності

### Фільтр Ходріка-Прескотта

Цей метод використовується для виділення трендового компонента у довгостроковому періоді. Для застосування методу слід вибрати змінну та виконати команду **Procs→Hodric-Prescott Filter...** У вікні (рис. 19) слід вказати назву нового ряду, що міститиме оцінку трендового компонента, і параметр згладжування.

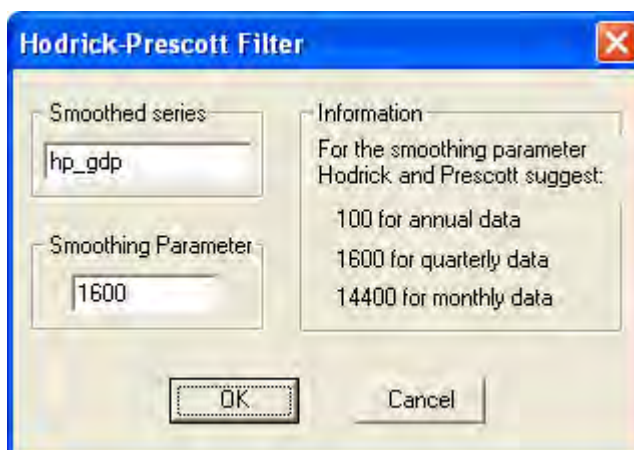


Рис. 19. Побудова методу Ходріка-Прескотта

В результаті виконання команди EViews створює нову змінну *hp\_gdp* та виводить графік, на якому розташовані вихідний та обчислений ряди (рис. 20).

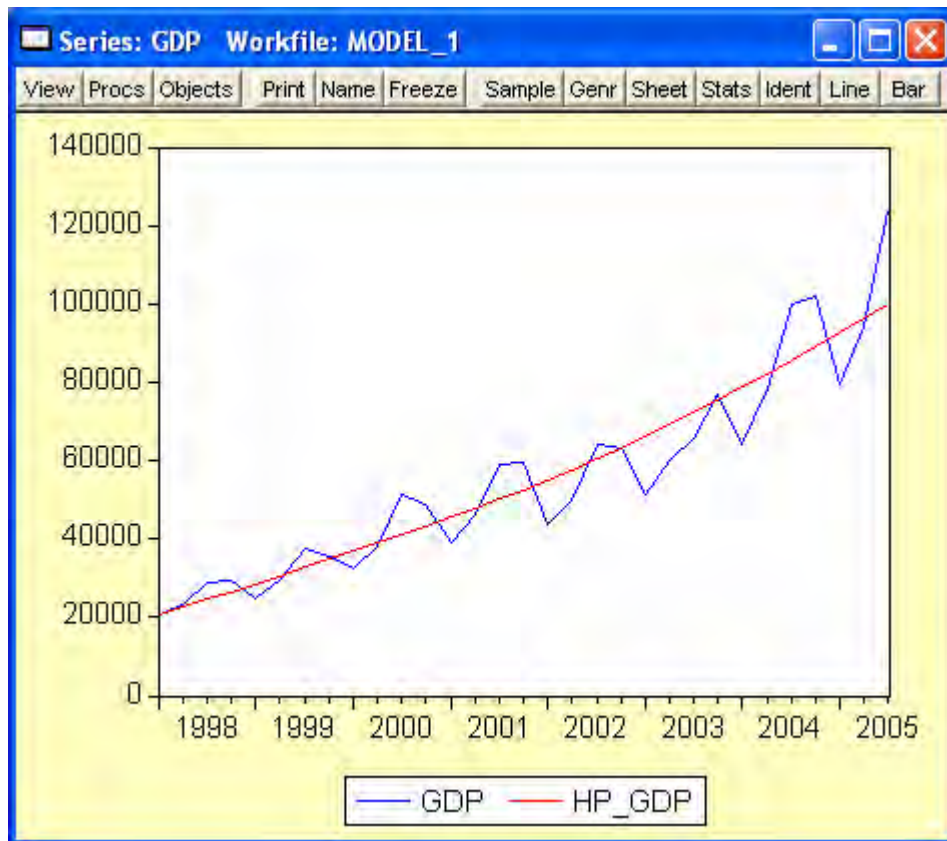


Рис. 20. Виділення тренду за допомогою фільтра Ходріка-Прескотта

### Лабораторна робота

**Задача 3.1.** Для кожного часового ряду в базі даних застосувати метод аналізу:

- метод усереднення (з кроками 3, 4, 5);
- експоненціальне згладжування;
- подвійне експоненціальне згладжування;
- потрійне експоненціальне згладжування;
- несезонну модель Холта-Вінтерса;
- адитивну модель Холта-Вінтерса;
- мультиплікативну модель Холта-Вінтерса;
- фільтр Ходріка-Прескотта.

Підрахувати прогнози на 2006 рік, скориставшись даними за 1996-2005 роки. Обчислити похибки прогнозування. Визначити найкращий метод виділення трендового компонента для кожного часового ряду.

**Задача 3.2.** Розробити інформаційну систему прогнозування на основі відомих Вам методів згладжування.

## РОЗДІЛ 4. ВИДІЛЕННЯ ТРЕНДОВОГО ТА СЕЗОННОГО КОМПОНЕНТІВ

### Теоретична частина

#### Функціональні види тренду

Стаціонарні процеси зустрічаються дуже рідко, але деякі часові ряди, що представляють економічну інформацію, часто можуть бути перетворені до стаціонарних, що значно спрощує їх аналіз. Перш за все, для часових рядів притаманна наявність трендового компонента. Його, як правило, моделюють як визначену функцію від часу, тому ми будемо представляти часовий ряд у вигляді  $y_t = f(t) + u_t$ , де  $f(t)$  представляє тренд часового ряду,  $u_t$  – стаціонарний часовий ряд. В залежності від часу ця величина може змінюватися, міняти напрям, тому іноді буває дуже важко точно встановити вигляд відповідної функції  $f$ .

Вибір виду тренду є досить складною задачею. Велику роль відіграє інтуїція дослідника, його наукова ерудиція і досвід. Слід бути дуже обережним при використанні для опису тренду функцій, що можуть різко змінювати напрямок росту (наприклад: поліном високого ступеня).

Типові класи “кривих зростання”, які застосовуються для моделювання економічних показників, наведені у таблиці.

#### Основні види функції тренду

Вид функції тренду	Опис
$f(t) = a_0 + a_1 t$	лінійний тренд
$f(t) = a_0 + a_1 t + a_2 t^2$	квадратичний тренд
$f(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$	поліноміальний тренд
$f(t) = a_0 e^{a_1 t}$	експоненціальний тренд
$f(t) = a_0 t^{a_1}$	показниковий тренд
$f(t) = \frac{a_0}{1 + a_1 t}$	гіперболічний тренд
$f(t) = a_0 + a_1 \ln(t)$	логарифмічний тренд
$f(t) = \frac{a_0}{1 + a_1 e^{a_2 t}}$	логістичний тренд
$f(t) = a_0 a_1^{a_2^t}$	тренд Гомперца
$f(t) = a_0 \left(1 - \frac{t - a_1}{a_2}\right)^{-a_3} \left(1 - \frac{t - a_4}{a_5}\right)^{-a_6}$	тренд Пірсона

## Моделювання сезонності

Часто економічна інформація залежить від того, в який період року спостерігається процес. Наприклад, споживання товарів збільшується у четвертому кварталі кожного року і зменшується у першому. Тому необхідно було розробити методи виділення таких сезонних коливань для більш точного прогнозування. Одним з варіантів є побудова регресійних моделей на основі фіктивних змінних.

**Фіктивні або бінарні змінні** (*dummy variables*) приймають тільки значення 0 або 1. Наприклад, якщо ми розглядаємо часовий ряд з квартальною структурою даних, то доцільним є розгляд такої моделі:

$$y_t = \beta_0 + \beta_1 q_1 + \beta_2 q_2 + \beta_3 q_3 + \varepsilon_t,$$

де  $q_1$  приймає значення 1, якщо  $t$  відповідає першому кварталу року, 0 - у інших випадках. Таким чином, вектор  $q_1$  має вигляд

$$q_1 = (1, 0, 0, 0, 1, 0, 0, 0, 1, 0, \dots)'$$

Тоді у перший квартал  $y_t = \beta_0 + \beta_1 q_1 + \varepsilon_t$ . В свою чергу  $q_2$  приймає значення 1, якщо  $t$  відповідає другому кварталу року, 0 - у інших випадках, тобто

$$q_2 = (0, 1, 0, 0, 0, 1, 0, 0, 0, 1, \dots)'$$

і  $y_t = \beta_0 + \beta_2 q_2 + \varepsilon_t$ . Нарешті вектор  $q_3$  має вигляд  $q_3 = (0, 0, 1, 0, 0, 0, 1, 0, 0, 0, \dots)'$ , а

$y_t = \beta_0 + \beta_3 q_3 + \varepsilon_t$ . Зауважимо, що для розгляду  $s$  сезонів достатньо мати  $s - 1$  сезонну змінну, оскільки рівню останнього  $s$ -го сезону відповідає константа:  $y_t = \beta_0 + \varepsilon_t$ .

Вищенаведена модель оцінюється за допомогою звичайного МНК, знаходяться коефіцієнти  $\beta_i$ ,  $i = \overline{0, 3}$ . Після цього можна робити прогноз на кожний з кварталів року.

Якщо розглядати часовий ряд з місячною структурою, то, звичайно, необхідно мати 11 фіктивних змінних.

Можливо, часовий ряд крім сезонного компонента має і трендовий. В такому разі, модель можна розширити:

$$y_t = \beta_0 + \alpha t + \beta_1 q_1 + \beta_2 q_2 + \beta_3 q_3 + \varepsilon_t.$$

Якщо в результаті аналізу залишків виявиться, що  $\hat{\varepsilon}_t$  не відповідають процесу білого шуму, то їх можна моделювати як деякий  $ARMA(p, q)$  процес.

## Приклад

Побудувати сезонну регресію для прибутків та збитків підприємств України, використовуючи дані за 1998-2000 роки.

**Розв'язок.** Оскільки дані наведені у місячній структурі, то можна побудувати дві моделі: одна на основі місячних даних, інша – на основі агрегованих квартальних. Оцінки коефіцієнтів моделей наведені у таблиці:

Коефіцієнт	Моделювання по місячних даних		Моделювання по квартальних даних	
	Прибуток	Збиток	Прибуток	Збиток
$\beta_0$	5212,5	8866,4	5712,5	8838,8
$\alpha$	50,2	20,3	470,6	201,7
$\beta_1$	-4038,1	-7759,9	-2281,9	-4852,2
$\beta_2$	-4555,4	-8303,8	-2700,1	-8294,7
$\beta_3$	-3329,1	-6356,7	-2181,9	-6969,3
$\beta_4$	-4435,0	-9022,9		
$\beta_5$	-4660,1	-9152,3		
$\beta_6$	-3404,8	-7843,7		
$\beta_7$	-4449,7	-8876,0		
$\beta_8$	-4355,6	-8596,9		
$\beta_9$	-3017,4	-7064,5		
$\beta_{10}$	-4619,0	-8726,0		
$\beta_{11}$	-5002,6	-8823,4		
$R^2$	0,797	0,904	0,901	0,848

За допомогою цих коефіцієнтів побудовано прогнози до кінця 2007 року.

### Прогноз прибутків та збитків підприємств України на 2001-2007 роки, млн. гривень

Рік	Моделювання по місячних даних		Моделювання по квартальних даних	
	Прогноз прибутків	Прогноз збитків	Прогноз прибутків	Прогноз збитків
2001	42265,7	26239,7	42978,3	26938,9
2002	49489,2	29167,6	50507,1	30166,4
2003	56712,6	32095,5	58036,0	33394,0
2004	63936,1	35023,3	65564,9	36621,5
2005	71159,6	37951,2	73093,8	39849,0
2006	78383,1	40879,1	80622,6	43076,5
2007	85606,5	43807,0	88151,5	46304,1

### Практична частина

#### Оцінка економетричних рівнянь

Економетричні рівняння є найбільш поширеною формою аналіз часових рядів. Побудова та оцінка рівнянь є дуже простою справою у EViews. Для створення нової моделі необхідно вибрати команду меню **Quick→Estimate**



**Equation....** У вікні слід вказати форму моделі, вибірку, по якій буде оцінена модель, та метод оцінки регресії. Опції дають можливість змінювати метод оцінки для врахування гетероскедастичності залишків або нестандартного вигляду регресії.

Для створення моделі слід перерахувати всі залежні і незалежні змінні. Наприклад, для побудови моделі виду

$$gdp_t = \beta_0 + \beta_1 gdp_{t-1} + \beta_2 inv_{t-1} + \varepsilon_t$$

слід вказати параметри, показані на рис. 21.

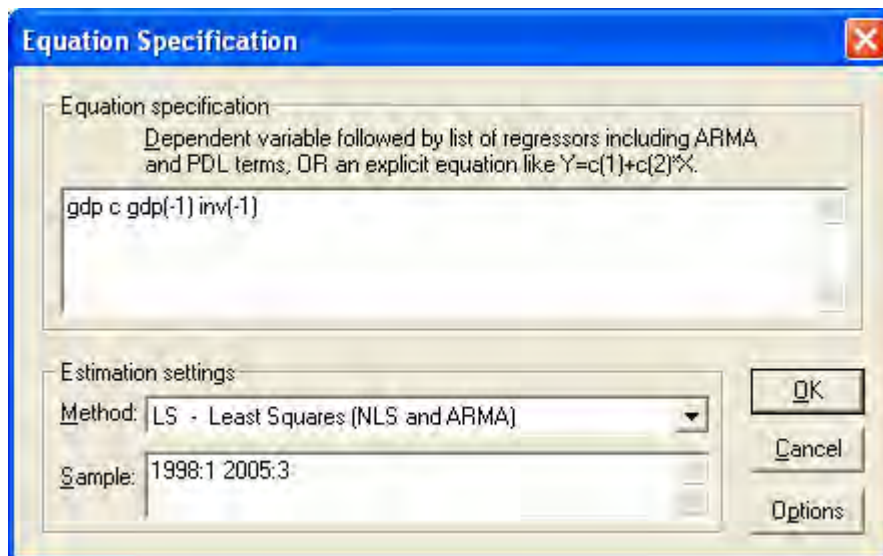


Рис. 21. Побудова регресії

Equation: UNTITLED Workfile: MODEL_1				
View Procs Objects Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: GDP				
Method: Least Squares				
Date: 01/13/06 Time: 14:16				
Sample(adjusted): 1998:2 2005:3				
Included observations: 30 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7455.713	4894.834	-1.523180	0.1393
GDP(-1)	1.554847	0.145784	10.66403	0.0000
INV(-1)	-2.021851	0.458417	-4.410069	0.0001
R-squared	0.884260	Mean dependent var	56669.57	
Adjusted R-squared	0.875686	S.D. dependent var	25223.74	
S.E. of regression	8893.425	Akaike info criterion	21.11865	
Sum squared resid	2.14E+09	Schwarz criterion	21.25877	
Log likelihood	-313.7798	F-statistic	103.1404	
Durbin-Watson stat	1.689045	Prob(F-statistic)	0.000000	

Рис. 22. Результати оцінки регресії.

У наведеному прикладі залежною змінною виступає  $gdp$ , незалежними змінними – лагові значення змінних  $gdp$  та  $inv$ . Параметр  $c$  використовується для використання у регресії константи.

Після відповідних обчислень EViews видає результати (рис. 22).

Верхня частина містить загальну інформацію по побудованій регресії (залежна змінна, метод оцінки, час оцінки, вибірка). Наступний блок містить інформацію про значення оцінок коефіцієнтів регресії, середньоквадратичні відхилення, значимість коефіцієнтів). Неважко побачити, що оцінки коефіцієнтів при змінних  $GDP(-1)$  та  $INV(-1)$  є значимими ( $Prob < 0.05$ ), а константа регресії не є значимою. Третій блок містить інформацію про якість побудованої регресії: коефіцієнт детермінації  $R^2 = 0,884$ , скоригований коефіцієнт детермінації  $R^2_{adj} = 0,876$ , середньоквадратичне відхилення регресії  $\sigma = 8893,425$ , сума квадратів залишків  $RSS = 2,14 \cdot 10^9$ , логарифм функції максимальної правдоподібності  $\ln L = -313,78$ , статистика Дарбіна-Уотсона  $DW = 1,69$ , вибіркове середнє  $\overline{gdp}$  та стандартне відхилення  $\sigma_{gdp} = 25223,74$  залежної змінної, інформаційний критерій Акайке  $AIC = 21,12$ , критерій Шварца  $SC = 21,26$ , практичне значення статистики Фішера для перевірки моделі на адекватність  $F_{pr} = 103,14$ . Значення  $Prob = 0,00 < 0,05$  свідчить про прийняття гіпотези про адекватність моделі.

Якщо є потреба змінити модель, то це можна зробити вибравши команду **Estimate** на панелі.

Звичайно, регресію можна представити у більш простому вигляді за допомогою меню **View→Representations** (рис.23). Повернутися до попереднього вигляду дозволяє меню **View→Estimation Output** або кнопка **Stats**.

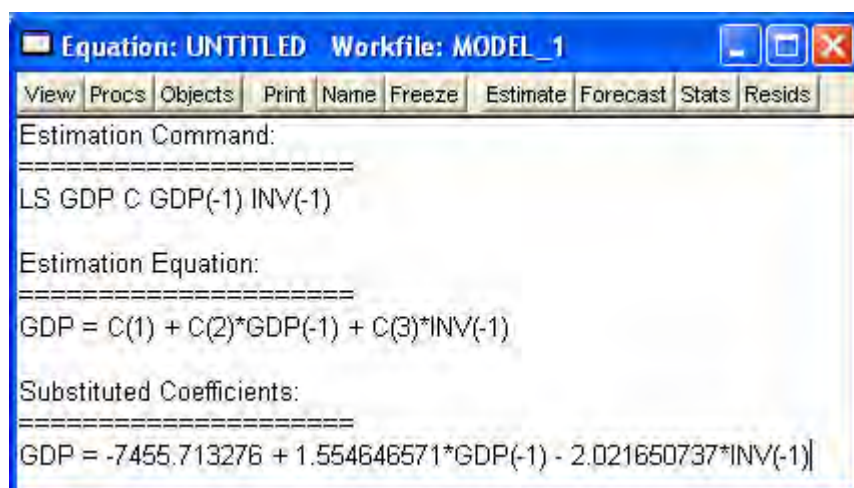


Рис. 23. Підстановка коефіцієнтів моделі до регресії.

Команда меню **View→Actual, Fitted, Residual→Actual, Fitted, Residual Table** дозволяє переглянути числові характеристики залежної змінної, її прогноз за регресією та залишки моделі. Команда **View→Actual, Fitted,**

**Residual→Actual, Fitted, Residual Graph** або кнопка **Resids** показує цю ж інформацію у графічній формі (рис. 24). Зверніть увагу, що залежна змінна та її оцінка відкладені по правій шкалі, а залишки моделі – по лівій.

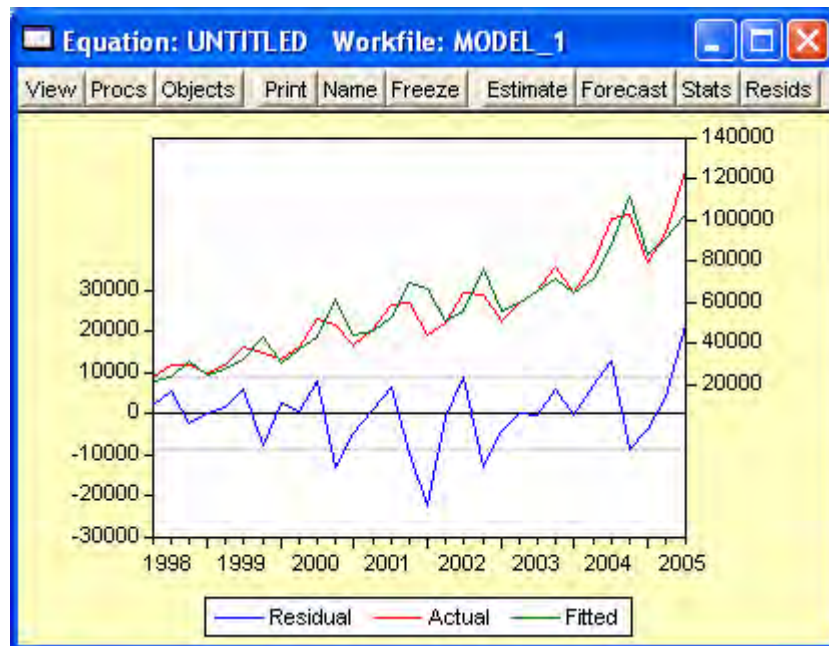


Рис. 24. Графічне представлення регресії.

Команда **View→Actual, Fitted, Residual→Residual Graph** виводить окремо графік залишків моделі, а команда **View→Actual, Fitted, Residual→Standardized Residual Graph** виводить графік залишків, нормованих на середньоквадратичне відхилення цих залишків.

Дослідити ступінь залежності між змінними допомагає коваріаційна матриця (**View→Covariance Matrix**).

Для збереження побудованої моделі в робочій книзі слід натиснути кнопку **Name** і вказати потрібне ім'я. При цьому створюється новий об'єкт зі значком

Ще однією можливістю моделі є здатність обраховувати прогнози. Зокрема, кнопка **Forecast** або меню **Procs→Forecast...** дозволяють відкрити вікно, в якому вказується:

- назва змінної, що міститиме прогнозні значення;
- вибірка, для якої слід обрахувати прогнози;
- метод прогнозування (динамічний, що враховує отримані прогнози для підрахунку наступних значень; статичний, що опирається лише на реальні значення часового ряду);
- опції (наприклад, вивід графіка, вивід надійних інтервалів).

При обрахунку прогнозів слід пам'ятати, що вибірка не може перевищувати діапазон значень, тому для розробки прогнозів для періодів, для яких не існує реальних даних, варто розширити діапазон за допомогою команди **Procs→Change Workfile Range....**

Як відомо, при оцінці будь-якої моделі залишки записуються у змінну *Resid*, проте існує можливість записати їх у іншу змінну за допомогою команди **Procs→Make Residual Series...**

При побудові моделі можуть використовуватися різноманітні змінні та функції системи.

### Моделювання сезонності

Одним з найпоширеніших способів моделювання сезонності є використання фіктивних змінних. Для побудови регресії з фіктивними змінними слід застосувати функцію системи *@seas(N)*. Як і всі функції системи її запис починається зі знаку „@”. Ця функція має один параметр *N* – номер сезону у році. Наприклад, для оцінки регресії виду

$$gdp_t = \beta_0 + \beta_1 q_1 + \beta_2 q_2 + \beta_3 q_3 + \varepsilon_t$$

у формі регресії (Quick→Estimate Equation...) використовується запис

$$gdp \ c \ @seas(1) \ @seas(2) \ @seas(3)$$

В наведеному прикладі використано 3 фіктивних змінні. Оскільки у робочій книзі дані мають квартальну структуру, то EViews створює фіктивні змінні саме з такою структурою. Якби дані мали місячну структуру, то необхідно було б використати 11 фіктивних змінних.

### Трендові моделі

EViews дозволяє легко оцінювати трендові моделі. Для створення трендової змінної, що містить послідовні цілі числа, використовується функція *@trend(N)*. Якщо параметр *N* не вказано, то першим елементом нової змінної буде 0, наступним 1, далі – 2, 3 тощо. Якщо ж в якості параметра вказана дата, то елемент змінної, що відповідає цій даті, буде дорівнювати 0, наступні елементи послідовно дорівнюватимуть 1, 2, 3 і т.д, а попередні – прийматимуть послідовні від’ємні значення: -1, -2, -3 і т.д.

Виділимо зі змінної *gdp* лінійний тренд, базою якого буде 2 квартал 2000 року. Запис відповідної регресії показано на рис. 25.

Аналогічним чином можна виділяти й інші види трендів. Наприклад, для виділення логарифмічного тренду слід записати таку форму моделі:

$$gdp \ c \ log(@trend)$$

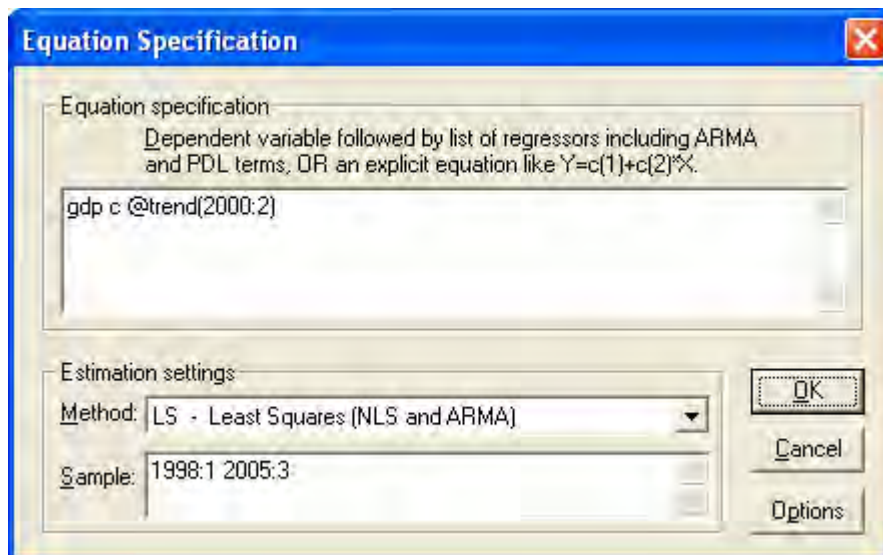


Рис. 25. Виділення лінійного тренду

### Перевірка гіпотези про значення коефіцієнтів моделі

Нехай для моделі

$$gdp_t = \beta_0 + \beta_1 gdp_{t-1} + \beta_2 inv_{t-1} + \varepsilon_t$$

слід перевірити гіпотезу

$$H_0 : \begin{cases} \beta_1 = 1.55, \\ \beta_2 = -2.2. \end{cases}$$

Перевірка гіпотези для побудованої моделі здійснюється за допомогою критерію Вальда (**View→Coefficient Tests→Wald Test (Coefficient Restrictions)...**). У діалоговому вікні вказується вид гіпотези (рис. 26). Зверніть увагу, що константа регресії має індекс 1, наступний коефіцієнт при  $gdp(-1)$  – 2, при  $inv(-1)$  – 3.

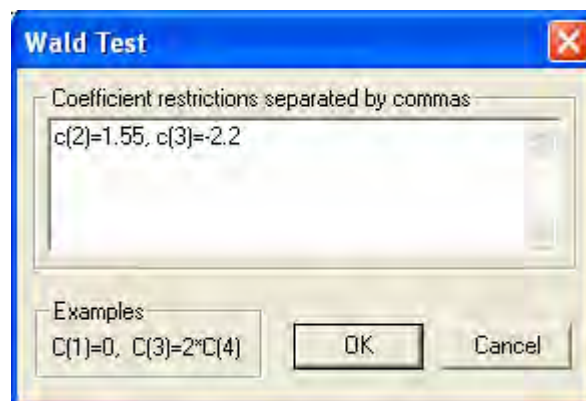


Рис. 26. Перевірка гіпотези про значення коефіцієнтів

Вивід результатів (рис. 27) свідчить, що нульова гіпотеза має бути прийнятою, оскільки значення *Probability* значно перевищує рівень 0.05.



Equation: EQ01    Workfile: MODEL_1			
View   Procs   Objects   Print   Name   Freeze   Estimate   Forecast   Stats   Res			
Wald Test: Equation: EQ01			
Null Hypothesis: C(2)=1.55 C(3)=-2.2			
F-statistic	0.346382	Probability	0.710339
Chi-square	0.692763	Probability	0.707243

Рис. 27. Вивід результатів перевірки гіпотези про значення коефіцієнтів

Нагадаємо, що значенням F–статистики можна користуватися лише за припущення про нормальний розподіл збурень регресії, в інших випадках варто використовувати значення  $\chi^2$ .

Аналогічно можна перевіряти довільні (навіть, нелінійні) залежності між коефіцієнтами регресії. Наприклад, для перевірки гіпотези

$$H_0 : \begin{cases} \beta_0 + \beta_1 = 1, \\ \beta_1 \beta_2 = \beta_0^2, \end{cases}$$

форму для тесту слід заповнити наступним чином:

$$c(1)+c(2)=1, c(2)*c(3)=c(1)^2$$

### Тестування про зайві та пропущені змінні

Для перевірки гіпотези про недоцільність включення додаткових змінних до регресії (**View→Coefficient Tests→Omitted Variables–Likelihood Ratio...**) у вікні вказують перелік додаткових змінних, які можна було б включити до моделі. Наприклад, розглянемо доцільність включення до моделі

$$gdp_t = \beta_0 + \beta_1 gdp_{t-1} + \beta_2 inv_{t-1} + \varepsilon_t$$

змінної *m2* (рис. 28).

Отримані значення F–статистики та максимуму функції правдоподібності свідчать, що змінну *m2* варто включити до моделі.

Аналогічно проводиться тестування щодо зайвих змінних (**View→Coefficient Tests→Redundant Variables–Likelihood Ratio...**). Наприклад, проведемо тестування щодо доцільності змінної *inv(-1)* у регресії (рис. 29).

Значення F–статистики та максимуму функції правдоподібності свідчать про важливість змінної *inv(-1)* у розглянутій регресії (*Probability*<0.05).

Equation: EQ01    Workfile: MODEL\_1

View   Procs   Objects   Print   Name   Freeze   Estimate   Forecast   Stats   Resids

Omitted Variables: M2

F-statistic	34.85721	Probability	0.000003
Log likelihood ratio	25.51302	Probability	0.000000

Test Equation:

Dependent Variable: GDP

Method: Least Squares

Date: 01/21/06    Time: 17:46

Sample: 1998:2 2005:3

Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8665.823	4252.777	2.037686	0.0519
GDP(-1)	0.765953	0.165150	4.637930	0.0001
INV(-1)	-1.668712	0.311139	-5.363239	0.0000
M2	0.363346	0.061542	5.904000	0.0000

R-squared	0.950552	Mean dependent var	56669.57
Adjusted R-squared	0.944847	S.D. dependent var	25223.74
S.E. of regression	5923.725	Akaike info criterion	20.33488
Sum squared resid	9.12E+08	Schwarz criterion	20.52171
Log likelihood	-301.0233	F-statistic	166.6027
Durbin-Watson stat	2.214708	Prob(F-statistic)	0.000000

Рис. 28. Перевірка гіпотези про доцільність включення до моделі додаткових змінних

### Лабораторна робота

**Задача 4.1.** Для Ваших часових рядів побудувати різні види трендів:

- лінійний тренд;
- квадратичний тренд;
- поліноміальний тренд;
- експоненціальний тренд;
- показниковий тренд;
- гіперболічний тренд;
- логарифмічний тренд;
- логістичний тренд.

Зробити на їх основі прогнози на поточний рік, підрахувати помилку прогнозування.

Equation: EQ01    Workfile: MODEL\_1

View   Procs   Objects   Print   Name   Freeze   Estimate   Forecast   Stats   Resids

Redundant Variables: INV(-1)

F-statistic	19.44871	Probability	0.000148
Log likelihood ratio	16.27535	Probability	0.000055

Test Equation:  
Dependent Variable: GDP  
Method: Least Squares  
Date: 01/21/06    Time: 18:07  
Sample: 1998:2 2005:3  
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3536.144	5425.899	0.651716	0.5199
GDP(-1)	0.998213	0.094060	10.61251	0.0000

R-squared	0.800889	Mean dependent var	56669.57
Adjusted R-squared	0.793778	S.D. dependent var	25223.74
S.E. of regression	11454.52	Akaike info criterion	21.59450
Sum squared resid	3.67E+09	Schwarz criterion	21.68791
Log likelihood	-321.9174	F-statistic	112.6253
Durbin-Watson stat	1.859213	Prob(F-statistic)	0.000000

Рис. 29. Перевірка гіпотези про доцільність виключення з моделі змінних

**Задача 4.2.** Побудувати для кожного часового ряду сезонну регресійну модель з трендом, обчислити прогнози, похибку прогнозування.

**Задача 4.3.** Розробити та перевірити гіпотези про значення коефіцієнтів регресій.

**Задача 4.4.** Визначити необхідність включення до побудованих регресій (зад. 4.1, 4.2) додаткових змінних.

**Задача 4.5.** Виділити з усіх змінних трендові компоненти. Побудувати найкращу економетричну модель з отриманих часових рядів.

**Задача 4.6.** Нехай часовий ряд місячних спостережень має лінійний тренд та сезонний компонент:  $y_t = a + bt + S_t + \varepsilon_t$ , причому  $S_t = S_{t-12}$  для всіх  $t$ ,  $\varepsilon_t$  – стаціонарний ряд випадкових збурень. Показати, що застосування оператора  $\Delta_{12}$  для часового ряду приводить до його стаціонарного. Чи буде оператор  $\Delta_{12}$  зводити часовий ряд до стаціонарного, якщо  $y_t = (a + bt)S_t + \varepsilon_t$ ?



### Теоретична частина

#### Процес «білого шуму»

Для будь-якого часового ряду можна створити деяку модель, яка б відображала розвиток процесу. Але яка б не була модель, неможливо прогнозувати значення часового ряду без похибки. Звичайно, розроблені методи, які намагаються зменшити цю похибку, але ніколи не можна бути впевненим у абсолютній точності прогнозів. Іншими словами, ніколи не було, нема і не буде моделі, яка б могла робити прогноз абсолютно точно. Таким чином, будь-яка модель повинна мати як доданок похибку, за величиною якої вимірюється точність моделі:

$$y_t = f(x_t) + \varepsilon_t,$$

де  $x_t$  - набір екзогенних змінних.

При прогнозуванні стаціонарних часових рядів припускається, що похибка  $\varepsilon_t$  у кожний період часу  $t$  має властивості:

- 1) нульове математичне сподівання, тобто  $\mu = E\varepsilon_t = 0$ ;
- 2) постійну дисперсію, тобто  $\gamma_0 = \text{var}(\varepsilon_t) = \sigma^2$ ;
- 3) некорельованість елементів  $\gamma_j = \text{cov}(\varepsilon_t, \varepsilon_{t-j}) = 0, \quad j \neq 0$ .

Будь-який процес, який задовольняє цим властивостям, називається «білим шумом».

#### Лаговий оператор

Особливе значення при роботі з часовими рядами має **лаговий оператор**. Зазвичай його позначають  $B$  (від англійського back-shift) або  $L$  (від німецького der Lagoperator). За його допомогою можна отримувати значення часового ряду як функції від його інших значень. В математичному плані застосування лагового оператора для значення  $y_t$  видає  $y_{t-1}$ , тобто

$$By_t = y_{t-1}.$$

Можна використовувати лаговий оператор декілька разів поспіль:

$$B(By_t) = By_{t-1} = y_{t-2}.$$

Якщо лаговий оператор використовується  $k > 0$  разів, то число  $k$  записується як показник степеня:

$$B^k(y_t) = y_{t-k}, \quad k > 0.$$

В той же час абсолютно прийнятною є застосування лагового оператора з від'ємним показником степеня:

$$B^{-k}(y_t) = y_{t+k}, \quad k > 0.$$

При цьому ми отримаємо майбутнє значення часового ряду.

Нарешті, якщо ми застосуємо лаговий оператор 0 разів, ми отримаємо те саме значення часового ряду:

$$B^0 y_t = y_t.$$

При роботі з лаговими операторами бажано пам'ятати наступну тотожність:

$$\frac{1}{1-B} = 1 + B + B^2 + B^3 + \dots$$

За допомогою лагового оператора можна записувати поліноми виду:

$$A_m(B) = a_0 + a_1 B + a_2 B^2 + \dots + a_m B^m.$$

Таким чином, лаговий оператор можна розглядати як деяку змінну, проводячи над нею стандартні математичні дії, наприклад:

$$A_3(B) \cdot y_t = (a_0 + a_1 B + a_2 B^2 + a_3 B^3) y_t = a_0 y_t + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3},$$

$$\begin{aligned} A_4(B) \cdot y_{t-3} &= (a_0 + a_1 B + a_2 B^2 + a_3 B^3 + a_4 B^4) y_{t-3} = \\ &= a_0 y_{t-3} + a_1 y_{t-4} + a_2 y_{t-5} + a_3 y_{t-6} + a_4 y_{t-7}. \end{aligned}$$

### МА(q)-процес

Нехай  $\varepsilon_t$  є процесом “білого шуму”,  $q$  - деяке натуральне число, тоді процес  $MA(q)$  будується за правилом:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} = \mu + \Theta_q(B) \varepsilon_t,$$

де  $\Theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ .

Розглянемо основні характеристики цього процесу.

1. Математичне сподівання:

$$\begin{aligned} E y_t &= E \mu + E \varepsilon_t + E(\theta_1 \varepsilon_{t-1}) + E(\theta_2 \varepsilon_{t-2}) + \dots + E(\theta_q \varepsilon_{t-q}) = \\ &= E \mu + E \varepsilon_t + \theta_1 E \varepsilon_{t-1} + \theta_2 E \varepsilon_{t-2} + \dots + \theta_q E \varepsilon_{t-q} = \mu. \end{aligned}$$

2. Дисперсія:

$$\begin{aligned} \gamma_0 &= E(y_t - \mu)^2 = E(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q})^2 = \\ &= E(\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 + \theta_2^2 \varepsilon_{t-2}^2 + \dots + \theta_q^2 \varepsilon_{t-q}^2 + 2\theta_1 \varepsilon_t \varepsilon_{t-1} + 2\theta_2 \varepsilon_t \varepsilon_{t-2} + \dots + \\ &\quad + 2\theta_q \varepsilon_t \varepsilon_{t-q} + 2\theta_1 \theta_2 \varepsilon_{t-1} \varepsilon_{t-2} + \dots + 2\theta_1 \theta_q \varepsilon_{t-1} \varepsilon_{t-q} + \dots + 2\theta_{q-1} \theta_q \varepsilon_{t-q+1} \varepsilon_{t-q}) = \\ &= \sigma^2 + \theta_1^2 \sigma^2 + \theta_2^2 \sigma^2 + \dots + \theta_q^2 \sigma^2 = \sigma^2 (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2). \end{aligned}$$

3. Коваріація  $j$ -го порядку ( $j = 1, 2, 3, \dots$ ):

$$\begin{aligned} \gamma_j &= E((y_t - \mu)(y_{t-j} - \mu)) = \\ &= E((\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q})(\varepsilon_{t-j} + \theta_1 \varepsilon_{t-1-j} + \theta_2 \varepsilon_{t-2-j} + \dots + \theta_q \varepsilon_{t-q-j})) = \\ &= E(\theta_j \varepsilon_{t-j}^2 + \theta_{j+1} \theta_1 \varepsilon_{t-j-1}^2 + \theta_{j+2} \theta_2 \varepsilon_{t-j-2}^2 + \dots + \theta_{j+q} \theta_q \varepsilon_{t-q}^2) = \end{aligned}$$

$$= \begin{cases} (\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_q\theta_{q+j})\sigma^2, & j = 1, 2, \dots, q, \\ 0, & j > q. \end{cases}$$

Звідси випливає, що якщо  $q$  є скінченним числом, то  $MA(q)$ –процес є **стаціонарним**.

Для того, щоб  $MA(\infty)$ –процес був також **стаціонарним**, необхідно, щоб дисперсія та коваріація були скінченними числами. Очевидно, це можливо за умови, якщо

$$\sum_{j=0}^{\infty} \theta_j^2 < \infty.$$

### Процес авторегресії $AR(p)$

Розвиток процесів авторегресії пов'язаний з необхідністю аналізу моделей, у яких одна змінна залежить від своїх попередніх значень. Побудова звичайних лінійних регресій не могла дати повне уявлення про напрями розвитку таких часових рядів. Внаслідок цього були запроваджені процеси авторегресії.

Нехай  $\varepsilon_t$  – “білий шум”,  $p$  – ціле невід’ємне число, то процес авторегресії порядку  $p$  записується у вигляді:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t = c + \Phi(B)y + \varepsilon_t,$$

де  $\Phi(B) = \varphi_1 B + \varphi_2 B^2 + \dots + \varphi_p B^p$ .

Іноді, для зручності розрахунків будемо використовувати такий запис

$$\tilde{\Phi}(B)y_t = c + \varepsilon_t,$$

де  $\tilde{\Phi}(B) = 1 - \Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$ .

Процес  $AR(p)$  є **стаціонарним**, коли всі корені  $z_i$  рівняння

$$1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p = 0$$

задовольняють умові:  $|z_i| > 1$ .

Математичне сподівання дорівнює  $\mu = \frac{c}{1 - \varphi_1 - \varphi_2 - \dots - \varphi_p}$ . Для

знаходження автоковаріації використовується наступна система:

$$\gamma_j = E\left((y_t - \mu)(y_{t-j} - \mu)\right) = \begin{cases} \varphi_1 \gamma_1 + \varphi_2 \gamma_2 + \dots + \varphi_p \gamma_p + \sigma^2, & j = 0, \\ \varphi_1 \gamma_{j-1} + \varphi_2 \gamma_{j-2} + \dots + \varphi_p \gamma_{j-p}, & j = 1, 2, \dots \end{cases}$$

Щоб знайти у явному вигляді значення  $\gamma_j$  потрібно розв’язати всі рівняння системи.

Для  $AR(p)$ –процесу існує єдине  $MA(\infty)$ -представлення, коефіцієнти якого знаходяться за рекурентною формулою ( $s > 0$ ):

$$\psi_s = \varphi_1 \psi_{s-1} + \varphi_2 \psi_{s-2} + \dots + \varphi_p \psi_{s-p}, \text{ де } \psi_0 = 1, \psi_j = 0, \quad j < 0.$$

Кожен скінченний  $MA(q)$ -процес може бути представлений у вигляді нескінченного  $AR(\infty)$ -процесу.

В загальному випадку,  $MA(q)$ -процес можна виразити у вигляді нескінченного  $AR(\infty)$ -процесу, коли всі  $z_i$  корені рівняння  $1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$  задовольняють умові  $|z_i| > 1$ . В такому випадку коефіцієнти знаходяться з системи:

$$\eta_1 = -\theta_1,$$

$$\eta_2 = -\theta_1 \eta_1 - \theta_2,$$

...

$$\eta_s = -(\theta_1 \eta_{s-1} + \theta_2 \eta_{s-2} + \dots + \theta_q \eta_{s-q}), \quad s > 0,$$

причому для формалізації  $\eta_0 = 1, \quad \eta_j = 0, \quad j < 0$ .

Слід зауважити, що якщо деякий  $AR(\infty)$  відповідає  $MA(p)$ -процесу, то він є стаціонарним.

### ARMA-процес

Після вивчення властивостей  $AR(p)$ - та  $MA(q)$ -процесів є природнім розробити сумісний їх процес. Таким процесом є  $ARMA(p, q)$ , який являє собою суму двох розглянутих вище процесів. Відмінністю є те, що  $y_t$  залежить не тільки від своїх попередніх значень, або тільки від збурень минулих періодів, а одночасно від лагових змінних та збурень. Загальним видом  $ARMA(p, q)$ -процесу є:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}.$$

Це можна записати коротше:

$$\tilde{\Phi}(B) y_t = \Theta(B) \varepsilon_t, \text{ де}$$

$$\tilde{\Phi}(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p,$$

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q.$$

$ARMA(p, q)$ -процес є **стаціонарним**, коли всі корені  $z_i$  рівняння

$$1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p = 0$$

задовольняють умові  $|z_i| > 1$ .

$ARMA(p, q)$ -процес можна **перетворити** у нескінченний  $AR(\infty)$ - або  $MA(\infty)$ -процес, коли всі корені  $z_i$  рівняння  $1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_p z^p = 0$  задовольняють умові  $|z_i| > 1$ .

$ARMA(p, q)$ -процес можна подати або як нескінченний  $AR(\infty)$ , або як нескінченний  $MA(\infty)$ -процес.

### Прогнозування на основі ARMA-моделей

Будь-яка модель створюється для аналізу інформації та прогнозування. Розглянемо процес прогнозування на основі  $ARMA$ -моделей. Будемо позначати прогнозне значення часового ряду на  $\tau$  періодів вперед, яке робиться в момент

$t$  через  $\hat{y}_{t+\tau|t}$ . Для створення прогнозного значення будемо використовувати всю наявну інформацію на період  $t$ . Наша мета – мінімізувати один з критеріїв помилок. Зазвичай використовується критерій  $MSE$  (mean squared error):

$$MSE = E\left(y_{t+\tau} - \hat{y}_{t+\tau|t}\right)^2.$$

Припустимо, що відомі значення збурень за перші  $T$  періодів. Тоді, для нескінченного  $MA(\infty)$ -процесу

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots$$

виконується рівність

$$y_{t+\tau} = \mu + \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t-1+\tau} + \theta_2 \varepsilon_{t-2+\tau} + \dots + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots$$

Оскільки ми припускали, що математичне сподівання  $\varepsilon_t$  є нульовим, то замість очікуваного прогнозу на наступні періоди ми можемо використовувати нульове значення. Таким чином, прогнозом на період  $t + \tau$  в момент  $t$  є

$$\hat{y}_{t+\tau|t} = E_t y_{t+\tau} = \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots$$

Враховуючи, що всі збурення мають однакову дисперсію, підрахуємо середню похибку такого прогнозу:

$$y_{t+\tau} - \hat{y}_{t+\tau|t} = \varepsilon_{t+\tau} + \theta_1 \varepsilon_{t-1+\tau} + \theta_2 \varepsilon_{t-2+\tau} + \dots + \theta_{\tau-1} \varepsilon_{t+1},$$

$$MSE = E\left(y_{t+\tau} - \hat{y}_{t+\tau|t}\right)^2 = \left(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_{\tau-1}^2\right) \sigma^2.$$

Для скінченного  $MA(q)$ -процесу прогнозом на період  $t + \tau$  у період  $t$  є:

$$\hat{y}_{t+\tau|t} = \begin{cases} \mu + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \dots + \theta_{\tau+q} \varepsilon_{t-q}, & \tau = \overline{1, q}, \\ \mu, & \tau > q. \end{cases}$$

Похибка підраховується за формулою

$$MSE = \begin{cases} \sigma^2, & \tau = 1, \\ \sigma^2 \left(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_{\tau-1}^2\right), & \tau = \overline{2, q}, \\ \sigma^2 \left(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_{q-1}^2\right), & \tau > q. \end{cases}$$

При розгляданні  $AR$ -процесів, слід пам'ятати, що його можна перетворити до нескінченного  $MA$ -процесу. Наприклад, для  $AR(1)$ -процесу маємо:

$$y_t = c + \varphi_1 y_{t-1} + \varepsilon_t,$$

$$\mu = E y_t = \frac{c}{1 - \varphi_1},$$

$$y_t - \mu = \varphi_1 (y_{t-1} - \mu) + \varepsilon_t = \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_1^2 \varepsilon_{t-2} + \dots$$

Таким чином, прогноз на період  $t + \tau$  у період  $t$  складається як нескінченна сума:

$$\hat{y}_{t+\tau|t} = \mu + \varphi_1^\tau \varepsilon_t + \varphi_1^{\tau+1} \varepsilon_{t-1} + \varphi_1^{\tau+2} \varepsilon_{t-2} + \dots$$

Відповідно, похибка складає

$$MSE = \begin{cases} \sigma^2, & \tau = 1, \\ \sigma^2 \left( 1 + \varphi_1^2 + \varphi_1^4 + \dots + \varphi_1^{2(\tau-1)} \right), & \tau > 1. \end{cases}$$

Якщо  $\tau \rightarrow \infty$ , то  $MSE \rightarrow \frac{\sigma^2}{1 - \varphi_1^2}$ .

### ARIMA-процеси

$ARIMA(p, d, q)$ -процеси є загальною моделлю прогнозування нестационарних часових рядів, які можуть бути приведені до стаціонарного виду за допомогою логарифмування або різниць. Ці процеси мають багато часткових випадків, які є дуже цікавими для аналізу.

Запис процесу має три параметри:  $p$  – кількість лагових змінних,  $q$  – кількість лагових збурень,  $d$  – кількість різниць.

Під різницею розуміється трансформація часового ряду за правилом

$$\tilde{y}_t = \Delta y_t = (1 - B) y_t = y_t - y_{t-1} = \beta_0 + \Theta(B) \varepsilon_t, \quad t = \overline{2, T}.$$

Іноді, значення початкового ряду спочатку логарифмують, тоді

$$\tilde{y}_t = \ln \frac{y_t}{y_{t-1}} = \ln y_t - \ln y_{t-1}, \quad t = \overline{2, T}.$$

В цьому випадку нові значення  $\tilde{y}_t$  відповідають рівню зростання процесу. Отриманий після цього стаціонарний процес відрізняється від початкового кількістю спостережень.

Процес називається **інтегрованим** порядку  $d$ , коли  $(1 - B)^d y_t$  є стаціонарним та перетворюваним процесом.

Якщо після вищенаведених перетворень процес стає стаціонарним, то

$$(1 - B) y_t = \beta_0 + \Theta(B) \varepsilon_t.$$

За допомогою елементарних перетворень отримуємо:

$$\frac{1}{\Theta(B)} (1 - B) y_t = \frac{1}{\Theta(B)} \cdot \beta_0 + \varepsilon_t,$$

або  $N(B)(1 - B) y_t = \tilde{\beta}_0 + \varepsilon_t$ , де  $\tilde{\beta}_0 = \frac{1}{\Theta(B)}$ ,

$$\tilde{N}(B) y_t = \tilde{\beta}_0 + \varepsilon_t, \text{ де } \tilde{N}(B) = N(B)(1 - B).$$

Оскільки ми припустили, що отриманий процес є стаціонарним, то всі корені рівняння  $\tilde{N}(B) = 0$  за абсолютною величиною повинні бути більшими за одиницю. Але очевидно, що рівняння  $N(B) = 0$  має один з коренів, рівний 1. В такому випадку кажуть, що процес має **одиничний корінь**. Дуже важливим є визначення того, чи є одиничний корінь для процесу. Тоді за допомогою застосування різниць його можна звести до стаціонарного.

Розглянемо властивості  $ARIMA$ -процесів.

Для моделі  $(1 - B) y_t = \beta_0 + \Theta(B) \varepsilon_t$  підрахуємо прогнозні значення.

$$\Delta \hat{y}_{t+\tau|t} = \beta_0 + \varepsilon_t + \theta_\tau \varepsilon_t + \theta_{\tau+1} \varepsilon_{t-1} + \theta_{\tau+2} \varepsilon_{t-2} + \dots$$

Для того, щоб отримати прогноз  $\hat{y}_{t+\tau|t}$ , необхідно зазначити, що

$$y_{t+\tau} = y_t + (y_{t+1} - y_t) + (y_{t+2} - y_{t+1}) + \dots + (y_{t+\tau} - y_{t+\tau-1}),$$

звідки

$$\begin{aligned} \hat{y}_{t+\tau|t} &= y_t + \Delta \hat{y}_{t+1|t} + \Delta \hat{y}_{t+2|t} + \dots + \Delta \hat{y}_{t+\tau|t} = \\ &= y_t + \tau \cdot \beta_0 + (\theta_\tau + \theta_{\tau-1} + \dots + \theta_1) \varepsilon_t + (\theta_{\tau+1} + \theta_\tau + \dots + \theta_2) \varepsilon_{t-1} + \\ &\quad + (\theta_{\tau+2} + \theta_{\tau+1} + \dots + \theta_3) \varepsilon_{t-2} + \dots \end{aligned}$$

Помилка прогнозування становитиме

$$\begin{aligned} y_{t+\tau} - \hat{y}_{t+\tau|t} &= (y_t + \Delta y_{t+1} + \Delta y_{t+2} + \dots + \Delta y_{t+\tau}) - \\ &\quad - (y_t + \Delta \hat{y}_{t+1|t} + \Delta \hat{y}_{t+2|t} + \dots + \Delta \hat{y}_{t+\tau|t}) \end{aligned}$$

Оскільки

$$\Delta y_t = \beta_0 + \Theta(B) \varepsilon_t = \beta_0 + \varepsilon_t + \tilde{\Theta}(B) \varepsilon_t, \text{ то}$$

$$\begin{aligned} y_{t+\tau} - \hat{y}_{t+\tau|t} &= \varepsilon_{t+\tau} + (1 + \theta_1) \varepsilon_{t+\tau-1} + (1 + \theta_1 + \theta_2) \varepsilon_{t+\tau-2} + \\ &\quad + \dots + (1 + \theta_1 + \theta_2 + \dots + \theta_{\tau-1}) \varepsilon_{t+1} \end{aligned}$$

Звідси ми отримуємо

$$\begin{aligned} MSE &= E(y_{t+\tau} - \hat{y}_{t+\tau|t})^2 = \sigma^2 (1 + (1 + \theta_1)^2 + \\ &\quad + (1 + \theta_1 + \theta_2)^2 + \dots + (1 + \theta_1 + \theta_2 + \dots + \theta_{\tau-1})^2) \end{aligned}$$

### “Випадкове блукання”

Частковим випадком *ARIMA* процесів є так зване “випадкове блукання”.

Його загальний вигляд

$$(1 - B) y_t = \beta_0 + \varepsilon_t.$$

Якщо в цій моделі  $\beta_0 = 0$ , то це “випадкове блукання” без напрямку, якщо  $\beta_0 \neq 0$ , то - “випадкове блукання” з напрямком.

Всі властивості “випадкового блукання” прямо впливають з властивостей *ARIMA*-процесів. Прогнозне значення має вигляд:

$$\hat{y}_{t+\tau|t} = y_t + \tau \cdot \beta_0.$$

Помилка прогнозування:

$$MSE = \tau \cdot \sigma^2.$$

### Тестування на наявність одиничного кореня

Для правильної ідентифікації моделі необхідно знати, чи можна представити модель виду  $\varphi(B) y_t = \varepsilon_t$  у вигляді  $(1 - B) \tilde{\varphi}(B) y_t = \varepsilon_t$ . Тобто треба виявити, чи має рівняння  $\varphi(z) = 0$  корінь  $z = 1$ . Якщо такий корінь існує, то значно спрощується аналіз часового ряду завдяки його перетворенню у *ARIMA*-процес.

Суть **тесту Діккі–Фуллера** (*Dickey-Fuller ADF-test*) полягає у тестуванні значимості коефіцієнта при лаговому значенні залежної змінної у допоміжній регресії, яка утворюється з початкової шляхом віднімання з обох частин  $y_{t-1}$ . Наприклад, для  $AR(1)$ -моделі  $y_t = \mu + \rho y_{t-1} + \varepsilon_t$  слід перевірити гіпотезу  $H_0 : \rho = 1$  проти альтернативної  $H_1 : \rho < 1$ . Тоді для регресії

$$\Delta y_t = \mu + \gamma y_{t-1} + \varepsilon_t$$

слід перевірити гіпотезу  $H_0 : \gamma = 0$  проти альтернативної  $H_0 : \gamma \neq 0$ . При перевірці гіпотези використовується не стандартна статистика Стюдента, а розподіл Маккіннона (MacKinnon). У загальному випадку можна тестувати гіпотезу, включаючи до допоміжної регресії константу, лінійний тренд.

Провести тестування на наявність одиничного кореня можна за допомогою **тесту Філіпса–Перрона** (Phillips–Perron). Значення відповідної статистики асимптотично наближається до статистики тесту Діккі–Фулера.

### Аналіз ARIMA-процесів

Джордж Бокс та Гвілім Дженкінс розвинули у своїй роботі метод аналізу часових рядів, які можна задати у вигляді *ARIMA*-процесів. За допомогою цієї методики можна знаходити всі невідомі параметри, включаючи кількість лагових змінних. Вибір відповідної моделі для часового ряду здійснюється за три етапи.

**Ідентифікація моделі.** Через порівняння функцій автокореляції та часткових автокореляційних функцій часового ряду з теоретичними функціями різних *ARIMA*-процесів обираються параметри моделі:  $p$  - кількість лагових змінних,  $q$  - кількість лагових збурень,  $d$  - кількість різниць.

Першим кроком ідентифікації є визначення, чи є поданий часовий ряд стаціонарним, чи необхідно враховувати сезонні коливання. Якщо процес здається нестационарним, то застосовують метод різниць, тобто зі значень початкового ряду утворюють новий ряд за правилом

$$\tilde{y}_t = \Delta y_t = y_t - y_{t-1}, \quad t = \overline{2, T}.$$

Іноді значення початкового ряду спочатку логарифмують, тоді

$$\tilde{y}_t = \ln \frac{y_t}{y_{t-1}} = \ln y_t - \ln y_{t-1}, \quad t = \overline{2, T}.$$

На практиці достатньо одного чи двох застосувань цього методу. Якщо після цього отриманий ряд здається стаціонарним, то величину  $d$  можна вважати визначеною. Якщо дуже важко визначити, чи є ряд стаціонарним, то аналізується корелограма та часткова автокореляційна функція.

Згадаємо, що корелограмою для часових рядів є графік коефіцієнта автокореляції в залежності від його порядку. Для “білого шуму” при великих  $n$  наближений розподіл кожного  $\gamma_k$  є нормальним з нульовим середнім та дисперсією  $\frac{1}{n}$ . Тому надійним інтервалом з рівнем надійності 95%, у якому



повинні лежати  $\gamma_k$ , щоб процес можна було віднести до стаціонарних, буде  $\left[-\frac{2}{\sqrt{T}}, \frac{2}{\sqrt{T}}\right]$ . На практиці більш важливо дослідити загальну поведінку коефіцієнтів автокореляції, тому потрібно, щоб усі  $\gamma_k$  для  $k > q$  попадали у надійний інтервал. Тоді можна говорити, що ряд розвивається як  $MA(q)$ -процес.

Тлумачення значень функції автокореляції та її поведінки не є легкою справою. Розглянемо декілька можливостей.

- 1) Випадковий процес. Якщо спостерігається абсолютно випадковий процес з великою кількістю даних, то  $\rho_j \approx 0$  при всіх ненульових значеннях  $j$ .
- 2) Короткотермінова залежність. Стаціонарні процеси часто-густо мають короткотермінову залежність між спостереженнями, що проявляється в тому, що декілька перших коефіцієнтів автокореляції є відносно великими в абсолютному значенні по відношенню до всіх інших, які наближаються до нуля. Якщо перші коефіцієнти є додатними, то присутня додатна автокореляція, якщо ж перші коефіцієнти автокореляції постійно змінюють свій знак, то спостерігається від'ємна автокореляція.
- 3) Нестационарні процеси. Часові ряди, які мають чіткий трендовий компонент, значення  $\rho_j$  не буде наближатися до нуля, навіть при досить великих  $j$ . Тому, як правило, підрахунок функції автокореляції застосовують тільки для стаціонарних часових рядів.
- 4) Сезонні коливання. При наявності сезонних коливань корелограма представлятиме, як правило, затухаючі хвилі однакової частоти. Якщо виділити сезонний компонент, то функція автокореляції покаже справжню залежність між спостереженнями у часі.

Часткова кореляційна функція використовується для визначення кількості лагових змінних у  $AR(p)$ -процесі. Очевидно, що з теоретичної точки зору, оцінка коефіцієнта автокореляції  $k$ -го порядку при  $k > p$  повинна дорівнювати нулю. Таким чином, якщо процес згенерований  $AR(p)$ -процесом, то на частковій корелограмі графік повинен спадати до нуля, тобто належати довірчому інтервалу, при  $k > p$ . Як і у попередньому випадку, використовується інтервал  $\left[-\frac{2}{\sqrt{T}}, \frac{2}{\sqrt{T}}\right]$ .

На практиці графіки автокореляції та часткової автокореляції не завжди є доволі ясними, тому іноді є досить важким завданням визначити параметри моделі. Якщо на діаграмах не можна побачити наближення до нуля, використовують змішані  $ARMA$ -процеси. Іноді оцінюють декілька можливих моделей і обирають найкращу за допомогою третього етапу – діагностики моделі.

**Оцінка моделі.** За допомогою відповідної методики, найчастіше при максимізації функції правдоподібності, знаходяться коефіцієнти моделі.

**Діагностика моделі.** За допомогою різноманітних критеріїв, які перевіряють всі властивості залишків, придатність моделі, встановлюється, чи адекватна модель відповідному процесу.

За допомогою **Q-тесту Льюнга-Бокса** (Ljung-Box) можна перевіряти гіпотезу про те, що автокореляція залишків відсутня до лагу  $k$ . При цьому

обраховується статистика  $Q = T(T+2) \sum_{j=1}^k \frac{\rho_j^2}{T-j}$ , яка порівнюється з

теоретичним значенням  $\chi^2_{m-p-q}$ -розподілу. Також за допомогою цього тесту можна перевірити гіпотезу про те, що залишки моделі відповідають процесу „білого шуму”.

Після того, як оцінені коефіцієнти моделі, треба визначити наскільки добре ця модель описує часовий ряд. Загальна характеристика критеріїв з їх властивостями подана у таблиці.

**Числові критерії адекватності моделей**

Назва критерію	Формула підрахунку	Бажаний екстремум
Коефіцієнт детермінації	$R^2 = 1 - \frac{\vec{\varepsilon}'\vec{\varepsilon}}{\sum_{i=1}^T (y_i - \bar{y})^2}$	1
Зважений коефіцієнт детермінації	$\bar{R}^2 = 1 - \frac{T-1}{T-(p+q)}(1-R^2)$	1
Прогнозний критерій	$FC = \frac{\vec{\varepsilon}'\vec{\varepsilon}}{T-(p+q)} \left( 1 + \frac{p+q}{T} \right)$	min
Інформаційний критерій	$AIC = \ln \left( \frac{\vec{\varepsilon}'\vec{\varepsilon}}{T} \right) + \frac{2(p+q)}{T}$	min
Інформаційний критерій Шварца-Ріссанена	$SIC = \ln \left( \frac{\vec{\varepsilon}'\vec{\varepsilon}}{T} \right) + \frac{p+q}{T} \ln T$	min
Критерій Ханнана-Квіна	$HQ = \ln \left( \frac{\vec{\varepsilon}'\vec{\varepsilon}}{T} \right) + \frac{c(p+q)}{T} \ln \ln T, \quad c \geq 2$	min

Застосування всіх критеріїв одночасно може призвести до різних рішень. Наприклад, коефіцієнт детермінації може призводити до збільшення параметрів  $p$  та  $q$ , тому цей критерій застосовувати для прийняття рішень не варто. Зважений коефіцієнт детермінації навпаки – призводить до зменшення числа доданків у  $ARMA$ -процесах. Інші критерії є більш сильними. Найбільш цікавими є останні два критерії, які є спроможними, тобто при  $T \rightarrow \infty$  повністю ідентифікують істинну модель. Критерій  $AIC$ , як правило, призводить до збільшення числа параметрів.

## Практична частина

### ARIMA – моделі

Побудова адекватної *ARIMA*–моделі можлива при правильній специфікації її параметрів. Як правило, така процедура основана на аналізі корелограми часового ряду. За її допомогою дослідники визначають можливі величини параметрів моделі. Припустимо, що необхідно оцінити модель *ARIMA*(2, 2, 1) відносно змінної *gdp*. В такому випадку в якості залежної змінної регресії вказують функцію *d*, яка знаходить прирости відповідного порядку заданої змінної. Запис

$$d(gdp)$$

означає, що залежною змінною буде величина  $\Delta gdp_t = gdp_t - gdp_{t-1}$ . Якщо операцію взяття різниці слід використати декілька разів, то у функції вказується другий параметр. Вираз

$$d(gdp, 2)$$

знаходить величину  $\Delta^2 gdp_t = \Delta(gdp_t - gdp_{t-1}) = gdp_t - 2gdp_{t-1} + gdp_{t-2}$ .

Для побудови регресії з *AR* чи *MA* компонентами використовуються функції *ar* та *ma*. Параметри цих визначають номер відповідного лагу, який слід включити до моделі. Для кінцевого запису моделі *ARIMA*(2, 1, 1) слід записати:

$$d(gdp, 2) \text{ c } ar(1) \text{ ar}(2) \text{ ma}(1)$$

Програма видає результат оцінки моделі (рис. 30). Крім стандартного регресійного виводу видається інформація про корені рівнянь, за допомогою яких можна визначати стаціонарність процесу та його перетворювальність. В нашому випадку абсолютне значення коренів близьке до 1, що вимагає проведення додаткових тестувань на наявність одиничного кореня.

При потребі до *ARIMA*–моделі можна додати сезонний компонент. Наприклад, для квартальних даних варто додати сезонність з кроком 4, що перетворює модель *ARIMA*(2, 2, 1)

$$\Delta^2 gdp_t = c + \phi_1 \Delta^2 gdp_{t-1} + \phi_2 \Delta^2 gdp_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

на модель

$$\begin{aligned} \Delta^2 gdp_t = c + \phi_1 \Delta^2 gdp_{t-1} + \phi_2 \Delta^2 gdp_{t-2} + \phi_4 \Delta^2 gdp_{t-4} \\ - \phi_4 \phi_1 \Delta^2 gdp_{t-5} - \phi_4 \phi_2 \Delta^2 gdp_{t-6} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \end{aligned}$$

Для оцінки подібної моделі необхідно записати:

$$d(gdp, 2) \text{ c } ar(1) \text{ ar}(2) \text{ sar}(4) \text{ ma}(1)$$

Результати оцінки представлені на рис. 31. Для включення сезонних доданків для шоків слід застосувати функцію *sma* з вказуванням необхідного номеру лагу.

Equation: EQ03 Workfile: MODEL_1				
View Procs Objects Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: D(GDP,2)				
Method: Least Squares				
Date: 01/21/06 Time: 16:53				
Sample(adjusted): 1999:1 2005:3				
Included observations: 27 after adjusting endpoints				
Convergence achieved after 80 iterations				
Backcast: OFF (Roots of MA process too large)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-40.54257	161.9153	-0.250394	0.8045
AR(1)	-0.044365	0.139099	-0.318949	0.7526
AR(2)	-0.894488	0.146605	-6.101337	0.0000
MA(1)	-1.614693	0.458512	-3.521594	0.0018
R-squared	0.895670	Mean dependent var	1090.259	
Adjusted R-squared	0.882061	S.D. dependent var	16141.86	
S.E. of regression	5543.466	Akaike info criterion	20.21458	
Sum squared resid	7.07E+08	Schwarz criterion	20.40656	
Log likelihood	-268.8968	F-statistic	65.81795	
Durbin-Watson stat	3.334055	Prob(F-statistic)	0.000000	
Inverted AR Roots	-.02+.95i	-.02-.95i		
Inverted MA Roots	1.61			
Estimated MA process is noninvertible				

Рис. 30. Оцінка моделі ARIMA(2, 1, 1)

Корелограма залишків останньої побудованої моделі (рис. 32) свідчить, що для всіх значень лагів від 1 до 16 приймається гіпотеза про відповідність залишків процесу білого шуму (всі значення  $Prob > 0.05$ ).

Побудуємо прогнози за отриманою моделлю на наступні 5 кварталів. Для цього змінимо діапазон даних на „1998:1-2006:4”, та використаємо у вікні оціненої моделі кнопку **Forecast** (рис. 33). У вікні обирається змінна, яку слід прогнозувати, в нашому випадку *gdp*. Також вказується новий ряд, що міститиме прогнозні величини (*gdpf*), період прогнозування (з четвертого кварталу 2005 року до 4 кварталу 2006 року), вибирається динамічний метод прогнозування.

### Тестування на одиничний корінь

Тестування на наявність одиничного кореня здійснюється при аналізі змінної за допомогою команди **View→Unit Root Test...** (рис. 34). EViews пропонує 2 види тестів: Діккі-Фуллера та Філіпса-Перрона. Також користувач вибирає, для якої саме послідовності проводити тестування: для рівнів часового ряду, для перших різниць, для других різниць. Останні можливості потрібні, якщо часовий ряд має декілька одиничних коренів.

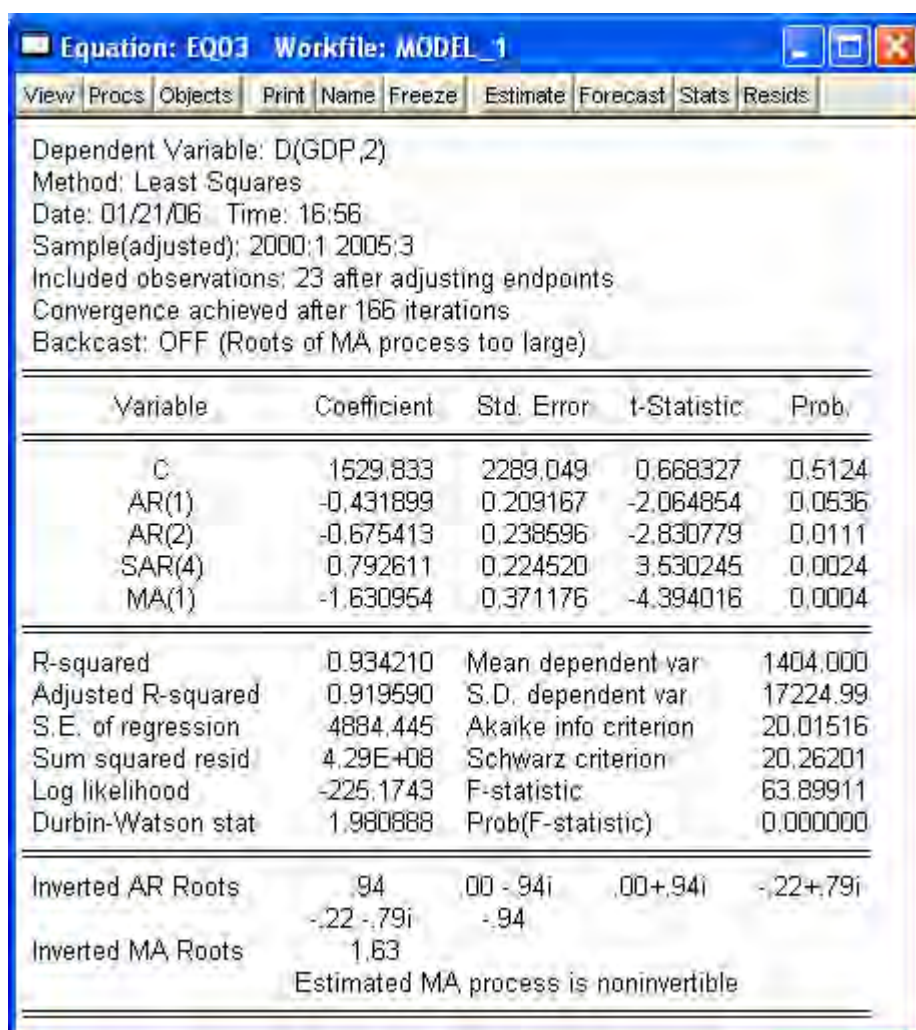


Рис. 31. Оцінка моделі ARIMA(2, 1, 1) з сезонними доданками

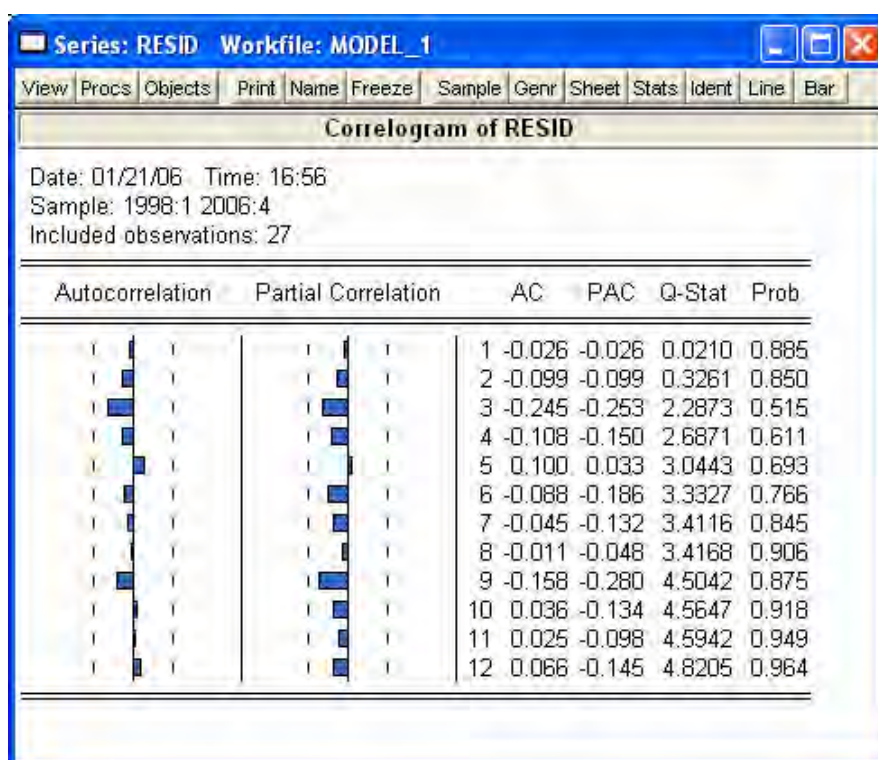


Рис. 32. Корелограма залишків оціненої моделі ARIMA(2, 1, 1) з сезонними доданками



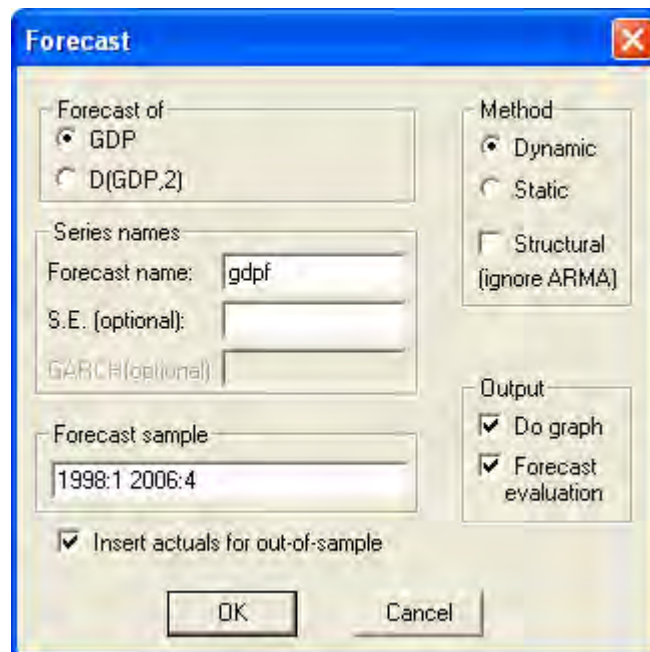


Рис. 33. Побудова прогнозу для змінної *gdp*

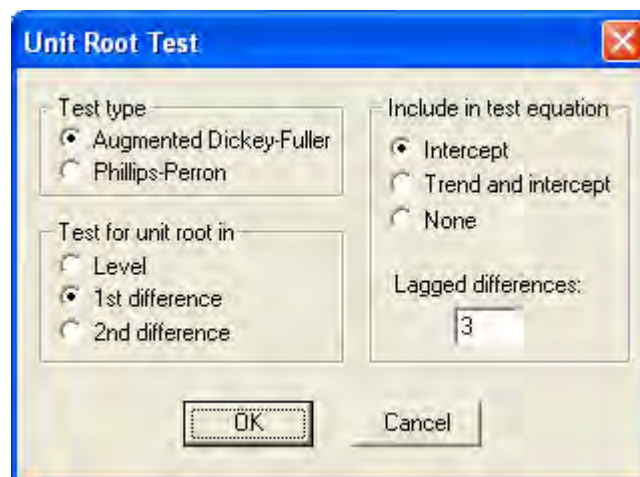


Рис. 34. Проведення тесту на наявність одиничного кореня

Далі користувач обирає змінні, що мають бути включені додатково до тестової моделі. EViews пропонує 3 можливості: включити лише константу, константу та лінійний тренд, не включати додаткові змінні.

Для тесту Діккі-Фуллера користувач вказує кількість лагових змінних, що включаються до тестової моделі. Наприклад, дані рис. 34 вказують на застосування наступної тестової моделі:

$$\Delta^2 gdp_t = \beta_0 + \beta_1 t + \beta_2 \Delta gdp_{t-1} + \beta_3 \Delta^2 gdp_{t-1} + \beta_4 \Delta^2 gdp_{t-2} + \beta_5 \Delta^2 gdp_{t-3} + \varepsilon_t.$$

Виконання тестової процедури (рис. 35) виводить статистику Діккі-Фулера (*ADF Test Statistic*) та критичні значення для різних рівнів надійності.

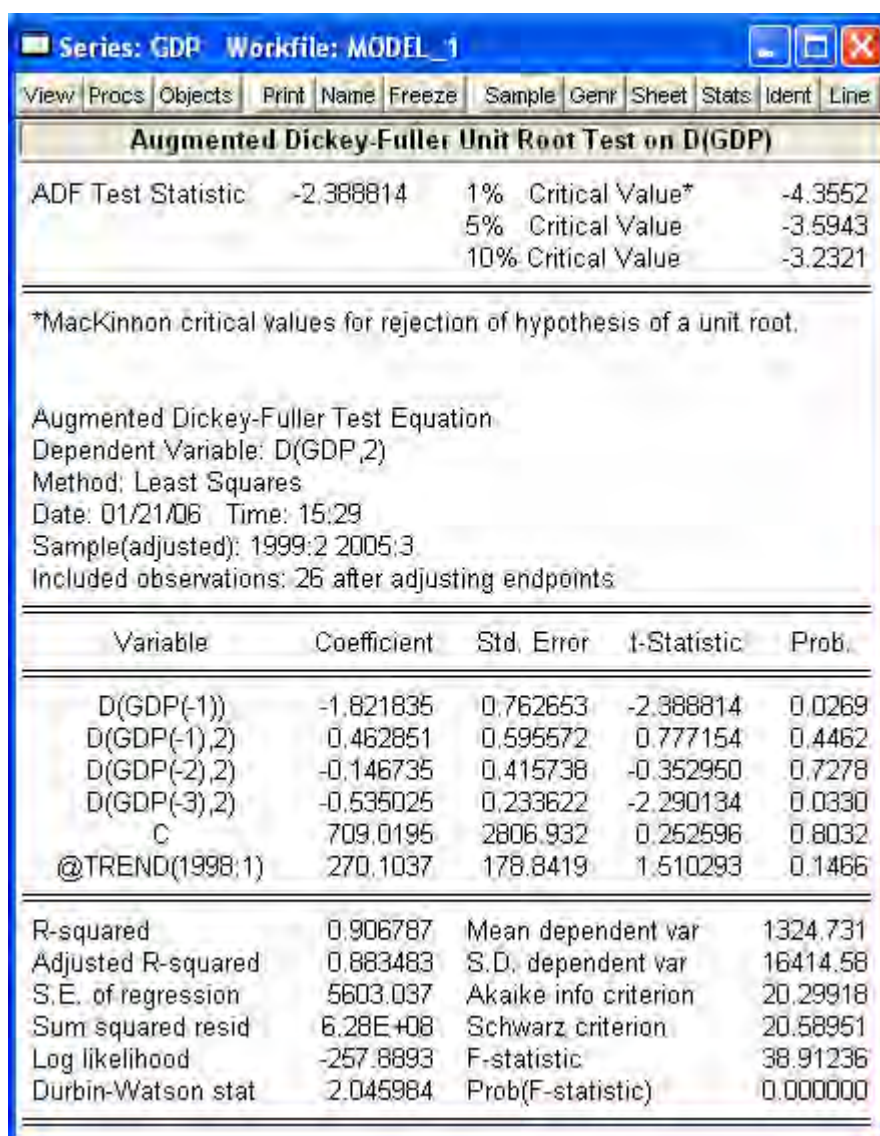


Рис. 35. Аналіз тесту Діккі-Фуллера на наявність одиничного кореня

Отримані дані свідчать, що неможливо відхилити гіпотезу про наявність одиничного кореня у часовому ряді  $d(gdp)$ . Якщо провести аналогічний тест для других різниць ряду, то гіпотеза про одиничний корінь буде відхилена для довільного рівня значимості. Таким чином, можна зробити висновок, що часовий ряд  $gdp$  є інтегрованим порядку 2, тобто  $gdp \sim I(2)$ .

При використанні тесту Філіпса-Перрона замість кількості включених до тестової моделі лагів вказується номер останнього лагу, який застосовується при обчисленні відповідної статистики.

### Інші тести

EViews має широкий набір стандартних тестів. Серед найбільш важливих ще не розглянутих вище є тест Уайта на гетероскедастичність збурень (**View**→**Residual Tests**→**White Heteroskedasticity**). Цей тест присутній в двох варіаціях: з попарними добутками (**cross terms**) та без них (**no cross terms**). В результаті виводиться значення  $F$ -статистики, за якого робиться висновок про

прийняття чи відхилення гіпотези. В нашому прикладі (рис. 36) гіпотеза про наявність гетероскедастичності відхиляється ( $Probability > 0.05$ ).

Equation: EQ01    Workfile: MODEL\_1

View   Procs   Objects   Print   Name   Freeze   Estimate   Forecast   Stats   Resids

White Heteroskedasticity Test:

F-statistic	2.114730	Probability	0.108884
Obs*R-squared	7.584453	Probability	0.108042

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 01/21/06    Time: 20:11

Sample: 1998:2 2005:3

Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-56618349	1.44E+08	-0.393951	0.6970
GDP(-1)	2553.042	6466.128	0.394833	0.6963
GDP(-1)^2	0.018479	0.050205	0.368078	0.7159
INV(-1)	-4512.561	15988.60	-0.282236	0.7801
INV(-1)^2	-0.207653	0.410529	-0.505819	0.6174

R-squared	0.252815	Mean dependent var	71183701
Adjusted R-squared	0.133266	S.D. dependent var	1.25E+08
S.E. of regression	1.16E+08	Akaike info criterion	40.12933
Sum squared resid	3.37E+17	Schwarz criterion	40.36286
Log likelihood	-596.9400	F-statistic	2.114730
Durbin-Watson stat	2.113814	Prob(F-statistic)	0.108884

Рис. 36. Тестування на наявність гетероскедастичності збурень

### Лабораторна робота

**Задача 5.1.** Виділити з часового ряду тренд за допомогою методу Ходріка-Прескотта. Провести нормування залишків за правилом:  $\tilde{y}_t = \frac{y_t - \bar{y}}{s.e.(y_t)}$ .

Перевірити новий ряд на відповідність процесу „білого шуму”.

**Задача 5.2.** Для отриманого у задачі 5.1 ряду за допомогою аналізу коредограми та графіку часткової кореляційної функції визначити параметри *ARMA*–моделі. Оцінити модель.

**Задача 5.3.** Побудувати та оцінити найкращу *ARIMA* модель для кожного часового ряду бази даних. Для визначення порядку інтегрованості процесу застосувати тестування на наявність одиничного кореня.

**Задача 5.4.** Для обчислених моделей (зад. 5.1–5.3) побудувати прогнози, обчислити похибки прогнозування.



**Задача 5.5.** Порівняти отримані моделі (зад. 5.1–5.3) на основі

- числових критеріїв;
- точності прогнозів;
- стабільності моделей.

**Задача 5.6.** Знайти параметри моделі  $AR(2)$ , якщо:

а).  $\rho_0 = 1, \rho_1 = \frac{13}{14}, \rho_2 = \frac{113}{140}$ ;

е).  $\rho_0 = 1, \rho_1 = -\frac{2}{3}, \rho_2 = 0$ ;

б).  $\rho_0 = 1, \rho_1 = -\frac{13}{14}, \rho_2 = \frac{113}{140}$ ;

ж).  $\rho_0 = 1, \rho_1 = \frac{1}{3}, \rho_2 = \frac{7}{15}$ ;

в).  $\rho_0 = 1, \rho_1 = \frac{11}{13}, \rho_2 = \frac{41}{65}$ ;

з).  $\rho_0 = 1, \rho_1 = -\frac{1}{3}, \rho_2 = \frac{7}{15}$ ;

г).  $\rho_0 = 1, \rho_1 = -\frac{11}{13}, \rho_2 = \frac{41}{65}$ ;

и).  $\rho_0 = 1, \rho_1 = \frac{1}{7}, \rho_2 = -\frac{13}{15}$ ;

д).  $\rho_0 = 1, \rho_1 = \frac{2}{3}, \rho_2 = 0$ ;

к).  $\rho_0 = 1, \rho_1 = -\frac{1}{7}, \rho_2 = -\frac{13}{15}$ .

**Задача 5.7.** Знайти параметри моделі  $ARMA(1, 1)$ , якщо

а).  $\rho_0 = 1; \rho_1 = -\frac{13}{19}; \rho_2 = -\frac{91}{190}$ ;

б).  $\rho_0 = 1; \rho_1 = \frac{11}{175}; \rho_2 = -\frac{11}{350}$ ;

в).  $\rho_0 = 1; \rho_1 = \frac{31}{41}; \rho_2 = \frac{93}{205}$ ;

г).  $\rho_0 = 1; \rho_1 = -\frac{38}{155}; \rho_2 = \frac{134}{775}$ ;

д).  $\rho_0 = 1; \rho_1 = \frac{38}{155}; \rho_2 = \frac{134}{775}$ ;

е).  $\rho_0 = 1; \rho_1 = \frac{31}{41}; \rho_2 = -\frac{93}{205}$ ;

ж).  $\rho_0 = 1; \rho_1 = \frac{31}{46}; \rho_2 = \frac{62}{230}$ ;

з).  $\rho_0 = 1; \rho_1 = \frac{19}{110}; \rho_2 = -\frac{19}{275}$ ;

и).  $\rho_0 = 1; \rho_1 = -\frac{19}{110}; \rho_2 = -\frac{19}{275}$ ;

к).  $\rho_0 = 1; \rho_1 = \frac{31}{46}; \rho_2 = -\frac{62}{230}$ .

## РОЗДІЛ 6. МОДЕЛІ ЗІ ЗМІННОЮ ДИСПЕРСІЄЮ

### Теоретична частина

Незважаючи на те, що за допомогою *ARIMA*-моделей можна описати доволі широкий клас часових рядів, іноді необхідно моделювати більш складні залежності, які можуть включати припущення про те, що залишки не є процесами “білого шуму”. Така ситуація трапляється при аналізі фінансових ринків, при вивченні курсів валют та акцій. Обумовлюється це тим, що на ринку одночасно діють багато агентів, які дотримуються однакової стратегії (наприклад, теорії оптимального портфелю), що призводить до доволі схожих рішень агентів. При цьому, ціна акцій або валют при використанні таких стратегій визначається як деяка функція від попередніх значень ціни. Оскільки на встановлення нової ринкової ціни впливають декілька великих груп агентів, кожна з яких пропонує свою ціну, то дійсна ринкова ціна стає функцією від функцій попередніх цін. Таким чином, спостерігається залежність між цінами, або, що є більш імовірним, між коливаннями ціни навколо деякого середнього рівня. Звідси, можна отримати такий результат, що залишки при прогнозуванні будуть залежати не тільки від рівня ринкової ціни або процесу, який її моделює, але й від попередніх залишків деяким чином. При цьому, дисперсія залишків є деякою функцією від своїх попередніх значень, тобто спостерігається гетероскедастичність залишків. Процеси, які дозволяють використовувати такі залежності, отримали назву *ARCH*-процесів (Autoregressive Conditional Heteroskedasticity). Існує декілька різновидів таких процесів.

### *ARCH*-моделі

Припустимо, що існує часовий ряд, з якого ми деяким чином виділили трендовий та сезонний компоненти. Таким чином, нехай  $\{Y_t\}$  – часовий ряд, який складається з залишків деякої регресії або процесу авторегресії. Це може також бути різницями деякого іншого ряду. Тоді його можна подати у вигляді

$$y_t = u_t,$$

де  $u_t$  – послідовність незалежних однаково розподілених випадкових величин з нульовим середнім та умовною дисперсією  $\sigma_t^2$ , тобто  $\sigma_t^2$  – це прогноз дисперсії на один період, оснований на попередніх значеннях. За визначенням, дисперсійне рівняння для *ARCH* (1)–процесу має вигляд:

$$\sigma_t^2 = \gamma + \delta u_{t-1}^2,$$

де  $\gamma, \delta$  вибрані таким чином, щоб  $\sigma_t^2 > 0$  для всіх  $t$ .

У більш загальному вигляді, *ARCH* ( $q$ )-процес залежить від  $q$  квадратів останніх лагових змінних:

$$y_t = u_t,$$

$$\sigma_t^2 = w + \sum_{j=1}^q \delta_j u_{t-j}^2.$$

При знаходженні коефіцієнтів  $ARCH(q)$ -моделі за допомогою методу найменших квадратів отримуються неефективні оцінки. Тому для оцінки коефіцієнтів всіх видів  $ARCH$ -моделей використовується метод максимальної правдоподібності. На відміну від звичайного методу, знаходження максимуму ускладнюється наявністю обмежень.

За допомогою  $ARCH(1)$ -моделі прогнози підраховуються за формулою:

$$E(y_{T+s}^2 | Y_T) = \alpha_0 (1 + \alpha_1 + \dots + \alpha_1^{s-1}) + \alpha_1^s y_T^2.$$

На основі  $ARCH$ -процесів, можна будувати інші моделі, які призначені для аналізу відповідних часових рядів.

### GARCH(p, q)-модель

Сама назва процесу (*Generalized Autoregressive Conditional Heteroskedasticity*) показує, що розглядається найбільш загальний вигляд модифікації  $ARCH$ -процесу. Розглянемо детально його структуру. Головне рівняння показує залежність часового ряду від деякого набору екзогенних змінних (ними можуть бути лагові значення часового ряду):

$$y_t = X_t \beta + u_t.$$

Дисперсійне рівняння містить середній рівень дисперсії  $w$ , інформацію про волатильність з попередніх періодів ( $u_{t-j}^2$  –  $ARCH$ -доданки), попередні значення дисперсії ( $\sigma_{t-j}^2$  –  $GARCH$ -доданки).

$$\sigma_t^2 = w + \delta_1 u_{t-1}^2 + \delta_2 u_{t-2}^2 + \dots + \delta_q u_{t-q}^2 + \theta_1 \sigma_{t-1}^2 + \theta_2 \sigma_{t-2}^2 + \dots + \theta_p \sigma_{t-p}^2 + \varepsilon_t.$$

Для збереження стаціонарності процесу, застосовуються такі обмеження:

$$w \geq 0, \delta \geq 0, \theta \geq 0, \sum_{i=1}^p \theta_i + \sum_{j=1}^q \delta_j < 1.$$

### TARCH – модель

Модель  $TARCH$  (*Threshold ARCH*) є модифікацією  $ARCH$ -процесу на випадок аналізу новин біржового ринку. Припустимо, що змінна  $d_t$  приймає значення 0, якщо новини „гарні”, тобто  $u_t > 0$ . В протилежному випадку, якщо  $u_t < 0$ , то  $d_t = 1$ .

Тоді  $TARCH$ -модель має вигляд:

$$y_t = X_t \beta + u_t$$

$$\begin{aligned} \sigma_t^2 = & w + \delta_1 u_{t-1}^2 + \delta_2 u_{t-2}^2 + \dots + \delta_q u_{t-q}^2 + \gamma u_{t-1}^2 d_{t-1} \\ & + \theta_1 \sigma_{t-1}^2 + \theta_2 \sigma_{t-2}^2 + \dots + \theta_p \sigma_{t-p}^2 + \varepsilon_t \end{aligned}$$

В цій моделі спостерігається різний ефект від новин. Якщо вплив гарних новин вимірюється коефіцієнтом  $\delta_1$ , то вплив поганих новин –  $\delta_1 + \gamma$ . Якщо  $\gamma$  статистично відмінний від 0, то вплив новин є несиметричним.

### EGARCH–модель

У моделі *EGARCH* логарифм дисперсії моделюється за допомогою формули

$$\ln \sigma_t^2 = w + \alpha \left| \frac{u_{t-1}}{\sigma_{t-1}} \right| + \beta \ln \sigma_{t-1}^2 + \gamma \frac{u_{t-1}}{\sigma_{t-1}}.$$

Такий вигляд дисперсійного рівняння гарантує невід'ємність умовної дисперсії. Ефект ліверджу показує коефіцієнт  $\gamma$ .

### Компонентна ARCH–модель

Якщо для *GARCH*(1,1) – процесу припустити, що існують різні короткотерміновий та довготерміновий середні рівні дисперсії, то дисперсійне рівняння можна подати у вигляді:

$$\sigma_t^2 = (1 - \delta - \theta)(1 - \rho)w + (\delta + \phi)u_{t-1}^2 - (\delta\rho + (\theta + \theta)\phi)u_{t-2}^2 + (\theta - \phi)\sigma_{t-1}^2 - (\theta\rho - (\delta + \theta)\phi)\sigma_{t-2}^2$$

Наведений запис свідчить, що компонентна *ARCH* модель еквівалентна нелінійному процесу *GARCH*(2, 2) без обмежень.

### Асиметрична компонентна ARCH–модель

Ця модель аналогічна попередній, однак до має і певні властивості *TARCH*–моделі. Загальний вигляд дисперсійного рівняння моделі:

$$\sigma_t^2 - q_t = \delta(u_{t-1}^2 - q_{t-1}) + \gamma(u_{t-1}^2 - q_{t-1})d_{t-1} + \theta(\sigma_{t-1}^2 - q_{t-1}) + \lambda_2 z_{2t},$$

де

$$q_t = w + \rho(q_{t-1} - w) + \phi(u_{t-1}^2 - \sigma_{t-1}^2) + \lambda_1 z_{1t},$$

$d_t$  – фіктивні змінні, що приймають значення 1, якщо спостерігається негативний шок  $u_t < 0$ ,

$z_1, z_2$  – екзогенні змінні.

### ARCH-M-процес

В цій моделі елемент дисперсійна складова додається прямо в головне рівняння:

$$y_t = X_t \beta + \phi \sigma_t^2 + u_t,$$

$$\sigma_t^2 = w + \delta_1 u_{t-1}^2 + \delta_2 u_{t-2}^2 + \dots + \delta_p u_{t-p}^2 + \theta_1 \sigma_{t-1}^2 + \theta_2 \sigma_{t-2}^2 + \dots + \theta_q \sigma_{t-q}^2 + \varepsilon_t.$$

Така модель використовується у фінансовій сфері при прогнозуванні, коли очікувана доходність активів пов'язана з очікуваним ризиком. Оцінений коефіцієнт  $\phi$  є мірою оберненої залежності між ризиком та доходністю.

## Практична частина

### Оцінка ARCH–процесів

Побудова GARCH-процесів аналогічна побудові регресій, проте у вікні для визначення форми моделі (**Quick→Estimate Equation...**) слід вибрати метод оцінки *ACRH* – *Autoregressive Conditional Heteroscedasticity*. При цьому вікно змінюється таким чином, щоб можна було вказати параметри відповідної моделі (рис. 37).

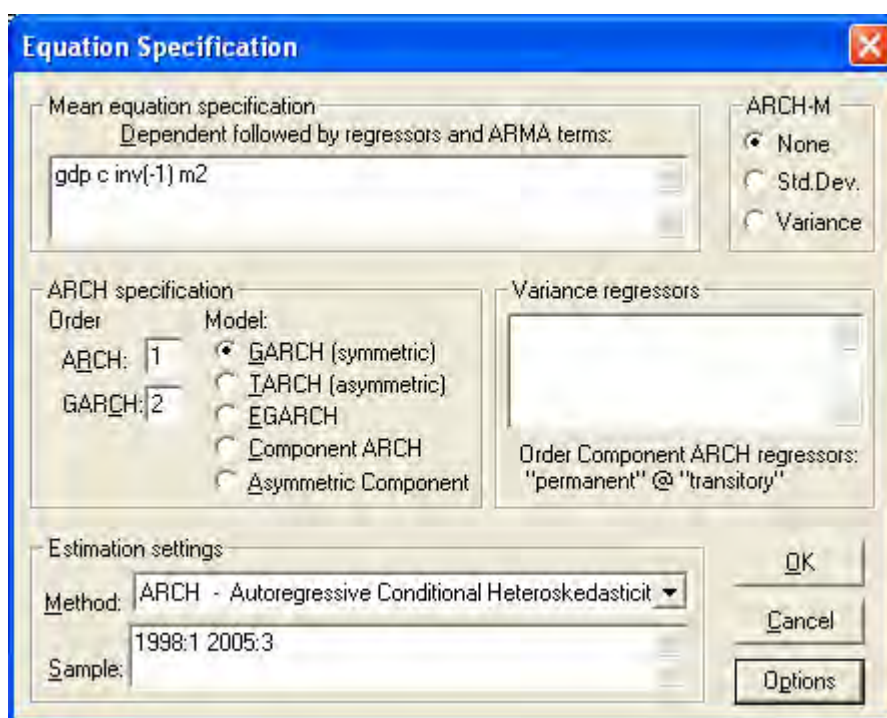


Рис. 37. Оцінка GARCH(1, 2)–процесу

У верхній лівій частині вказується вигляд головного рівняння. Запис аналогічний побудові звичайної регресії: спочатку вказується залежна змінна, потім – незалежні змінні. У лівій середній частині вказуються параметри  $p$  та  $q$  процесу, а також необхідна модель. Якщо є потреба дисперсійне рівняння можна розширити за допомогою розділу *Variance regressors*, в якому вказуються назви необхідних змінних. В разі необхідності розгляду *ARCH-M* моделі у розділі *ARCH-M* вказується змінна, що буде присутня у головному рівнянні: середньоквадратичне відхилення чи дисперсія.

Вивід результатів оцінки відрізняється від виводу при побудові звичайної регресії (рис. 38). Верхня та нижня частини містить інформацію про оцінки головного рівняння, а середня – про оцінку дисперсійного рівняння.

Equation: UNTITLED    Workfile: MODEL_1				
View   Procs   Objects   Print   Name   Freeze   Estimate   Forecast   Stats   Resids				
Dependent Variable: GDP				
Method: ML - ARCH (Marquardt)				
Date: 01/21/06    Time: 20:50				
Sample(adjusted): 1998:2 2005:3				
Included observations: 30 after adjusting endpoints				
Convergence achieved after 157 iterations				
Variance backcast: ON				
	Coefficient	Std. Error	z-Statistic	Prob.
C	25827.07	945.9679	27.30227	0.0000
INV(1)	-0.326189	0.234176	-1.392926	0.1636
M2	0.540052	0.048185	11.20786	0.0000
Variance Equation				
C	35813157	1.20E+08	0.298509	0.7653
ARCH(1)	0.462785	0.481792	0.960548	0.3368
GARCH(1)	-0.706224	1.183926	-0.596510	0.5508
GARCH(2)	0.594984	1.243332	0.478540	0.6323
R-squared	0.903905	Mean dependent var	58669.57	
Adjusted R-squared	0.878836	S.D. dependent var	25223.74	
S.E. of regression	8780.030	Akaike info criterion	20.91336	
Sum squared resid	1.77E+09	Schwarz criterion	21.24030	
Log likelihood	-306.7004	F-statistic	36.05756	
Durbin-Watson stat	1.542035	Prob(F-statistic)	0.000000	

Рис. 38. Вивід результатів оцінки GARCH(1, 2)–процесу

### Лабораторна робота

**Задача 6.1.** Після виділення трендового та сезонного компонентів з часових рядів обчисліть коефіцієнти моделей:

- $ARCH(1)$ ;
- $ARCH(2)$ ;
- $ARCH(3)$ ;
- $ARCH(4)$ .

За допомогою числових критеріїв адекватності оберіть найкращу модель. Підрахуйте прогнози на наступні періоди.

**Задача 6.2.** Для Ваших часових рядів оцініть моделі:

- $GARCH(1,1)$ ;
- $GARCH(1,2)$ ;
- $GARCH(2,1)$ ;
- $GARCH(2,2)$ .

В якості екзогенних змінних використайте (при необхідності) фіктивні змінні для виділення сезонних коливань, змінну часу для виділення трендового компонента.

За допомогою числових критеріїв адекватності оберіть найкращу модель. Підрахуйте прогнози на наступні періоди.

**Задача 6.3.** Оцініть для Ваших часових рядів відповідну *TARCH*–модель. Підрахуйте за її допомогою прогнози на наступні періоди.

**Задача 6.4.** Оцініть для Ваших часових рядів відповідну *EGARCH*–модель. Підрахуйте за її допомогою прогнози на наступні періоди.

**Задача 6.5.** Оцініть для Ваших часових рядів *ARCH–M*–модель. Підрахуйте за її допомогою прогнози на наступні періоди.

**Задача 6.6.** Оцініть для Ваших часових рядів відповідну компонентну *ARCH*–модель. Підрахуйте за її допомогою прогнози на наступні періоди.

**Задача 6.7.** Оцініть для Ваших часових рядів асиметричну компонентну *ARCH*–модель. Підрахуйте за її допомогою прогнози на наступні періоди.

**Задача 6.8.** Виберіть для кожного часового ряду найкращу модель за допомогою порівняння числових критеріїв адекватності.

**Задача 6.9.** Виділіть з Ваших часових рядів трендовий та сезонний компонент. Виконайте зад. 6.3–6.8 для отриманих рядів. Порівняйте якість нових моделей з раніше обчисленими. Зробіть прогнози за кращими моделями.

**Задача 6.10.** За допомогою числових критеріїв адекватності порівняйте моделі, отримані у зад. 6.3–6.9.

**Задача 6.11.** Знайдіть автокореляційну функцію для *ARCH*(1)–процесу.

**Задача 6.12.** За допомогою методу максимальної правдоподібності знайдіть оцінки коефіцієнтів *ARCH*(1)–процесу. Якими будуть прогнози за *ARCH*(1)–процесом?

**Задача 6.13.** Визначте математичне сподівання та дисперсію процесів:

- *ARCH*(1);
- *ARCH*(2);
- *ARCH*(3);
- *ARCH*(4);
- *GARCH*(1,1).

## Теоретична частина

На основі розглянутих вище моделей можна перейти до побудови систем моделей для одночасного аналізу декількох змінних. Ці змінні можуть залежати одна від іншої, а також, від лагових значень всіх змінних. За допомогою системи можна не тільки здійснювати прогнозування всіх змінних, але й аналізувати взаємозалежності між змінними, точно встановлювати їх структуру, вивчати вплив різноманітних зовнішніх факторів.

Системи, які складаються лише зі змінних, які залежать тільки від їх лагових значень та інших змінних, отримали назву *VAR (vector autoregressive)* моделей. Якщо при аналізі моделі використовуються лагові значення деякого процесу з властивостями “білого шуму”, то такі моделі називаються *VARMA (vector autoregressive moving average)*. Крім того, якщо замість значень часового ряду беруться послідовні різниці, то система називається *VARIMA (vector autoregressive integrated moving average)*. Якщо додати до стандартної *VARMA*-моделі декілька екзогенних змінних, то отримана модель називатиметься *VARMAX (vector autoregressive moving average with exogenous variables)*.

## Визначення VAR-моделі

Нехай досліджується  $n$  змінних, кожна з яких спостерігалася протягом  $T$  періодів. Тоді найпростіша *VAR* модель описується системою рівнянь

$$y_{i,t} = c_i + a_{i1}^{(1)} y_{1,t-1} + a_{i2}^{(1)} y_{2,t-1} + \dots + a_{in}^{(1)} y_{n,t-1} + \\ + a_{i1}^{(2)} y_{1,t-2} + a_{i2}^{(2)} y_{2,t-2} + \dots + a_{in}^{(2)} y_{n,t-2} + \dots + \\ + a_{i1}^{(p)} y_{1,t-p} + a_{i2}^{(p)} y_{2,t-p} + \dots + a_{in}^{(p)} y_{n,t-p} + \varepsilon_{i,t}, \quad i = \overline{1, n}.$$

У цій моделі введені такі позначення:

$i$  – номер змінної,

$a_{i,j}^{(k)}$  – коефіцієнти моделі,

$\varepsilon_i$  – векторні процеси “білого шуму”.

Для спрощення запису введемо нові позначення:

$$Y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \dots \\ y_{n,t} \end{pmatrix}, \quad C = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix}, \quad A_j = \begin{pmatrix} a_{1,1}^{(j)} & a_{1,2}^{(j)} & \dots & a_{1,n}^{(j)} \\ a_{2,1}^{(j)} & a_{2,2}^{(j)} & \dots & a_{2,n}^{(j)} \\ \dots & \dots & \dots & \dots \\ a_{n,1}^{(j)} & a_{n,2}^{(j)} & \dots & a_{n,n}^{(j)} \end{pmatrix}, \quad j = 1, \dots, p, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \dots \\ \varepsilon_{nt} \end{pmatrix}.$$

Тоді модель приймає вигляд

$$I \cdot Y_t = C + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t,$$



де  $I$  – одинична матриця розміру  $(n + 1) \times (n + 1)$ .

Використовуючи поліноміальний запис

$$A(L) = I - A_1 L - A_2 L^2 - \dots - A_p L^p, \quad \text{отримуємо запис VAR-моделі у}$$

матричній формі:

$$A(L)Y_t = C + \varepsilon_t.$$

Для аналізу моделі використовуються такі припущення:

1) математичні сподівання всіх  $\varepsilon_t$  дорівнюють нулеві.

2) всі процеси “білого шуму”  $\varepsilon_t$  незалежні між собою.

VAR-модель є **стаціонарною**, коли всі корені рівняння  $\det(I - A_1 z - A_2 z^2 - \dots - A_p z^p) = 0$  за абсолютною величиною є більшими за одиницю. Проте визначення того, чи є деяка VAR-модель стаціонарною, є досить складним завданням.

Оцінка моделі здійснюється за допомогою  $n$  застосувань методу найменших квадратів або за допомогою методу максимальної правдоподібності.

Для вибору оптимальної кількості лагів у системі використовуються декілька критеріїв: критерій порівняння функції максимальної правдоподібності ( $LR$ ), модифікований критерій Акайки ( $AIC$ ), критерій Шварца ( $SC$ ), Ханнана-Квіна ( $HQ$ ).

При застосуванні **критерію порівняння функції максимальної правдоподібності ( $LR$ )** спочатку визначаються параметри моделі та підраховується визначник коваріаційної матриці похибок  $D_1$  при  $p_1$  лагах. Далі зменшується кількість лагів до  $p_2$  і знаходиться відповідне значення  $D_2$ . Після цього обраховується статистика

$$\chi^2 = (T - p_1) \ln \frac{D_1}{D_2}.$$

Якщо  $\chi^2 < \chi^2_{w^2(p_1-p_2)}$ , то обираємо модель з меншим числом лагів  $p_2$ .

Достатньо простим критерієм вибору є  **$AIC$  критерій**, який для VAR( $p$ )-моделі модифікується до вигляду:

$$AIC(p) = \ln |\hat{V}| + \frac{2pn^2}{T},$$

де  $|\hat{V}|$  – визначник коваріаційної матриці залишків. Значенням  $p_{\max}$  обирається те, при якому мінімізується  $AIC(p)$ .

Прогнозування на основі VAR-моделей проходить абсолютно аналогічно прогнозуванню AR моделей. Для моделі VAR( $p$ ):

$$I \cdot Y_{t+\tau} = C + A_1 Y_{t+\tau-1} + A_2 Y_{t+\tau-2} + \dots + A_p Y_{t+\tau-p} + \varepsilon_{t+\tau}.$$

за допомогою закону ітеративних сподівань, отримаємо:

$$\begin{aligned}\hat{Y}_{t+1|t} &= C + A_1 Y_t + A_2 Y_{t-1} + \dots + A_p Y_{t+1-p}, \\ \hat{Y}_{t+2|t} &= C + A_1 \hat{Y}_{t+1|t} + A_2 Y_t + \dots + A_p Y_{t+2-p}, \\ &\dots \\ \hat{Y}_{t+\tau|t} &= C + A_1 \hat{Y}_{t+\tau-1|t} + A_2 \hat{Y}_{t+\tau-2|t} + \dots + A_p Y_{t+\tau-p}.\end{aligned}$$

### Причинність за Гренджером

При побудові VAR-моделі необхідно знати, чи потрібно взагалі використовувати всю наявну інформацію. Іноді буває, що побудова моделі з обмеженою кількістю інформації є більш вдалою. Причинність за Гренджером – це тест, що дозволяє визначити, чи впливає наявність однієї зі змінних на точність моделі чи ні.

Розглянемо дві змінні  $y_1 = (y_{1,t}, y_{1,t-1}, \dots, y_{1,1})$  та  $y_2 = (y_{2,t}, y_{2,t-1}, \dots, y_{2,1})$ . Будемо називати змінну  $y_2$  причиною за Гренджером для змінної  $y_1$ , якщо нерівність

$$\begin{aligned}MSE(\hat{y}_{1,t+\tau|t} | y_{1,t}, y_{1,t-1}, \dots, y_{1,1}, y_{2,t}, y_{2,t-1}, \dots, y_{2,1}) < \\ < MSE(\hat{y}_{1,t+\tau|t} | y_{1,t}, y_{1,t-1}, \dots, y_{1,1})\end{aligned}$$

виконується принаймні для одного  $\tau$ . Якщо ж

$$\begin{aligned}MSE(\hat{y}_{1,t+\tau|t} | y_{1,t}, y_{1,t-1}, \dots, y_{1,1}, y_{2,t}, y_{2,t-1}, \dots, y_{2,1}) \\ = MSE(\hat{y}_{1,t+\tau|t} | y_{1,t}, y_{1,t-1}, \dots, y_{1,1})\end{aligned}$$

для всіх  $\tau$ , то  $y_2$  не є причиною за Гренджером для змінної  $y_1$ .

Таким чином для нас важливо знати, чи може покращити нашу оцінку присутність додаткової інформації  $y_2$ . За Гренджером слід побудувати регресію виду

$$\begin{aligned}y_{1,t} &= c_1 + \alpha_1 y_{1,t-1} + \alpha_2 y_{2,t-2} + \dots + \alpha_p y_{1,t-p} + \\ &+ \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \dots + \beta_p y_{2,t-p} + \varepsilon_{1,t}\end{aligned}$$

та перевірити гіпотезу  $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ .

### Імпульсний аналіз

Розглянемо стандартну VMA-модель:

$$Y_t = M + H_0 \varepsilon_t + H_1 \varepsilon_{t-1} + H_2 \varepsilon_{t-2} + \dots, \text{ де } H_0 - \text{одинична матриця.}$$

Кожна матриця коефіцієнтів має вигляд

$$B_\tau = \begin{pmatrix} b_{11}^{(\tau)} & b_{12}^{(\tau)} & \dots & b_{1n}^{(\tau)} \\ b_{21}^{(\tau)} & b_{22}^{(\tau)} & \dots & b_{2n}^{(\tau)} \\ \dots & \dots & \dots & \dots \\ b_{n1}^{(\tau)} & b_{n2}^{(\tau)} & \dots & b_{nn}^{(\tau)} \end{pmatrix}.$$

Елемент цієї матриці  $b_{ij}^{(\tau)}$  показує, як зміниться значення  $y_i$  в залежності від  $j$ -го шоку  $\tau$  періодів назад. Таким чином

$$b_{ij}^{(\tau)} = \frac{\partial y_{i, t+\tau}}{\partial \varepsilon_{j, t}} = \frac{\partial y_{i, t}}{\partial \varepsilon_{j, t-\tau}}.$$

Вираз  $b_{ij}^{(\tau)}$  як функція від  $\tau$  називається **імпульсно-відповідною функцією**. За її допомогою можна досліджувати, який вплив на майбутні значення мають відповідні шоки в минулому.

### Коінтеграція

У стандартній VAR-моделі немає обмежень щодо взаємного руху декількох часових рядів. Коінтеграція дозволяє досліднику вводити до відповідної системи необхідний зв'язок між змінними, що призводить до більш точного оцінювання моделі.

Якщо для двох нестационарних часових рядів  $y_{1,t}$  та  $y_{2,t}$  існує деяка їх лінійна комбінація  $y_{1,t} - ky_{2,t}$ , яка є стаціонарною, то змінні  $y_{1,t}$  та  $y_{2,t}$  є **коінтегрованими**.

Нехай часові ряди  $y_{1,t}$  та  $y_{2,t}$  є інтегрованими порядку  $d$ , тобто  $I(d)$ . Тоді, як правило, лінійна комбінація цих двох рядів також буде  $I(d)$ . Якщо, існує лінійна комбінація цих рядів  $I(d-b)$ , тоді ці ряди називаються **коінтегрованими порядку  $(d, b)$** , що позначають  $CI(d, b)$ .

### Тест Йохансена на коінтегрованість

Першим кроком тесту є оцінка моделі

$$Y_t = \{\alpha + \beta t\} + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t,$$

де фігурні дужки позначають елементи, які можуть не входити до рівняння. Після цього знаходиться матриця коефіцієнтів  $\Pi = I - A_1 - A_2 - \dots - A_p$ . Нульова гіпотеза полягає в тому, що ранг матриці  $\Pi$  не перевищує деякого числа  $k - H_0: \text{rank} \Pi < k$ , а альтернативна гіпотеза –  $H_1: \text{rank} \Pi = k$ .

Статистика для перевірки гіпотези (*trace statistic*) має вигляд:

$-T \sum_{i=r+1}^k \ln(1 - \tilde{\lambda}_i)$ , де  $\tilde{\lambda}_i$  – оцінка максимальної правдоподібності  $i$ -го кореня рівняння Йохансена (всі корені мають бути впорядковані у порядку спадання).


Аналогічно перевіряється гіпотеза  $H_0: \text{rank} \Pi < k$  проти альтернативної  $H_1: \text{rank} \Pi = r + 1$ . В цьому випадку підраховується максимальна статистика (*max statistic*)  $-T \ln(1 - \tilde{\lambda}_{r+1})$ .

В обох випадках знайдені статистики порівнюються з критичними значеннями, розрахованими Йохансеном та Юселиусом (Juselius), та удосконаленими Остервальдом (Osterwald) та Ленумом (Lenum).

Таким чином, послідовно перевіряються гіпотези для всіх значень  $r$ , поки не буде знайдено відповідний рівень коінтеграції.

## Практична частина

### Оцінка систем регресійних рівнянь

Створення нової системи здійснюється за допомогою меню **Objects**→**New Object...**→**System**. У створеному вікні слід записати формули необхідних рівнянь. При записі системи до робочої книги створюється об'єкт зі значком .

Вигляд системи фактично є довільним. Користувач має право записувати різноманітні функції від змінних, коефіцієнтів системи, обмеження на коефіцієнти системи, різноманітні допоміжні функції тощо. В якості прикладу розглянемо систему:

$$\begin{cases} gdp_t = \beta_0 + \beta_1 gdp_{t-1} + \beta_2 m2 + \beta_3 inv_{t-1} + \varepsilon_t, \\ inv_t = \gamma_0 + \gamma_1 gdp_{t-1} + \gamma_2 inv_{t-1} + \gamma_3 inv + v_t, \\ v_t = \rho v_{t-1} + \varepsilon_t. \end{cases}$$

Запис цієї системи у EViews має вигляд, показаний на рис. 39.

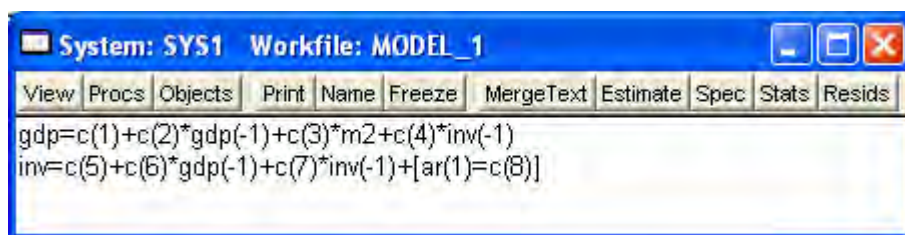


Рис. 39. Запис системи рівнянь

Слід пам'ятати, що при використанні різних методів необхідно вказувати, як правило, додаткову інформацію у правій частині вікна. Двоетапний МНК, триетапний МНК та узагальнені методи моментів використовуються для оцінки інструментальних змінних. Для цього слід записати при формуванні моделі список інструментальних змінних в кінці кожного рядка після службового знаку „@” (рис. 40).

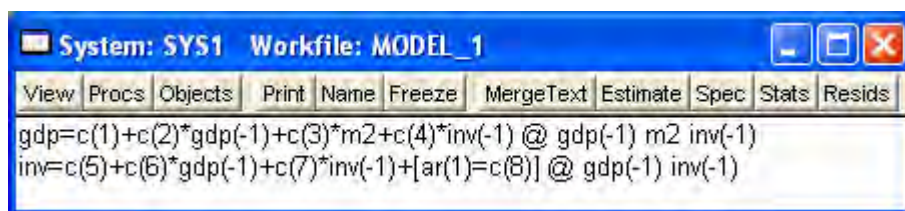


Рис. 40. Запис системи рівнянь з інструментальними змінними

Для оцінки моделі необхідно натиснути кнопку **Estimate** та вибрати один з дев'яти запропонованих методів оцінки (рис. 41).

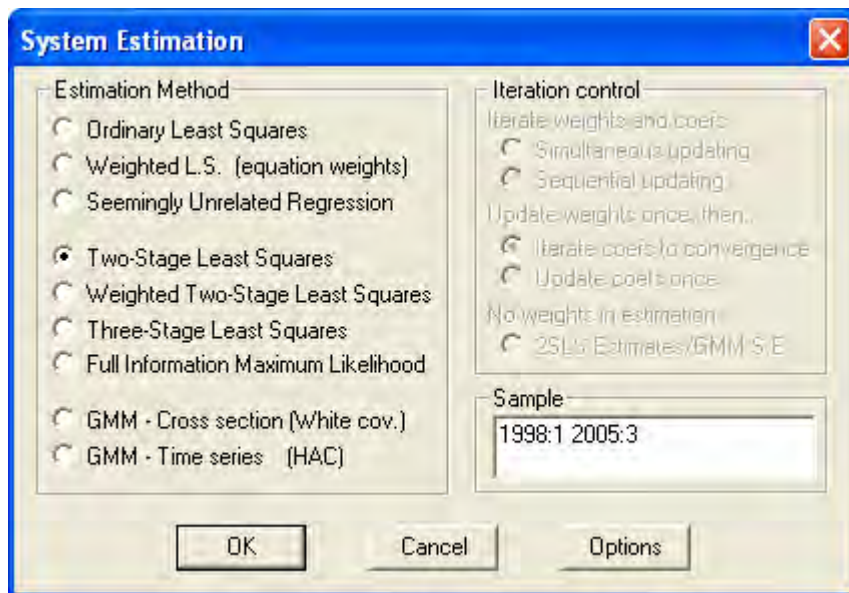


Рис. 41. Методи оцінки систем рівнянь

В результаті оцінки виводяться значення всіх коефіцієнтів моделі, а також інформація про якість кожного побудованого рівняння (рис. 42).

Після оцінки рівнянь користувач має можливість за допомогою меню **View** переглядати коваріаційні, кореляційні матриці коефіцієнтів, змінних моделі, проводити тестування гіпотези про обмеження на коефіцієнти системи, виводити графіки залишків кожного з рівнянь, будувати графіки рівнянь системи.

Майже аналогічно до систем реалізовані об'єкти типу модель. Такі об'єкти (**M**) являють собою набір різноманітних взаємопов'язаних рівнянь.

### VAR та VEC моделі

Для створення моделей *VAR*, *SVAR* та *VEC*, об'єкти яких позначаються значком **var**, використовується меню **Quick→Estimate VAR...** В лівому верхньому куті вибирається тип моделі: для оцінки стандартної *VAR* моделі слід обрати *Unrestricted VAR* (у цьому разі залишаються недоступними дві закладки вікна), для оцінки *VAR*–моделі з корекцією похибки, то обирається тип *Vector Error Correction*. Нижче вказується вибірка, за якою слід оцінювати модель.

У правому верхньому куті вказуються ендогенні змінні системи, нижче вказуються лаги ендогенних змінних, які слід включити до правих частин рівнянь. Слід зауважити, що номери лагів вказуються парами, наприклад, запис *1 3 6 6 8 12* означає, що слід включити лаги з першого по третій, шостий та з восьмого по дванадцятий. У нижньому правому куті вікна перераховуються екзогенні змінні системи (рис. 43).

В результаті обчислень EViews виведе оцінку двох рівнянь. Під кожним коефіцієнтом моделі будуть записані середньоквадратичні похибки та відповідні значення *t*–статистик (рис. 44).

System: SYS1    Workfile: MODEL\_1

View | Procs | Objects | Print | Name | Freeze | MergeText | Estimate | Spec | Stats | Resids

System: SYS1  
Estimation Method: Iterative Two-Stage Least Squares  
Date: 01/22/06    Time: 16:41  
Sample: 1998:2 2005:3  
Included observations: 30  
Total system (unbalanced) observations 59  
Convergence achieved after 7 iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	8665.823	4252.777	2.037686	0.0468
C(2)	0.765953	0.165150	4.637930	0.0000
C(3)	0.363346	0.061542	5.904000	0.0000
C(4)	-1.668712	0.311139	-5.363239	0.0000
C(5)	-11150.63	3040.366	-3.667527	0.0006
C(6)	0.568343	0.069893	8.131623	0.0000
C(7)	-1.034633	0.168582	-6.137282	0.0000
C(8)	0.419753	0.209454	2.004034	0.0504

Determinant residual covariance      2.77E+14

Equation:  $GDP=C(1)+C(2)*GDP(-1)+C(3)*M2+C(4)*INV(-1)$   
Observations: 30

R-squared	0.950552	Mean dependent var	56669.57
Adjusted R-squared	0.944847	S.D. dependent var	25223.74
S.E. of regression	5923.725	Sum squared resid	9.12E+08
Durbin-Watson stat	2.214708		

Equation:  $INV=C(5)+C(6)*GDP(-1)+C(7)*INV(-1)+[AR(1)=C(8)]$   
Observations: 29

R-squared	0.818168	Mean dependent var	10100.66
Adjusted R-squared	0.796348	S.D. dependent var	7357.885
S.E. of regression	3320.450	Sum squared resid	2.76E+08
Durbin-Watson stat	1.943153		

Рис. 42. Результати оцінки системи рівнянь за двоетапним МНК

Нижче виводиться інформація про якість кожного побудованого рівняння та системи в цілому (рис. 45).



**VAR Specification**

Basics | Cointegration | VEC Restrictions

VAR Type

☒ Unrestricted VAR

☐ Vector Error Correction

Endogenous Variables

gdp inv

Lag Intervals for Endogenous:

1 4

Estimation Sample

1998:1 2005:3

Exogenous Variables

c

OK Отмена

Рис. 43. Побудова стандартної VAR-моделі

Var: UNTITLED Workfile: MODEL\_1

View Procs Objects Print Name Freeze Estimate Stats

**Vector Autoregression Estimates**

Vector Autoregression Estimates  
Date: 01/22/06 Time: 17:11  
Sample(adjusted): 1999:1 2005:3  
Included observations: 27 after adjusting endpoints  
Standard errors in ( ) & t-statistics in [ ]

	GDP	INV
GDP(-1)	0.424341 (0.21285) [ 1.99358]	0.146388 (0.10702) [ 1.36788]
GDP(-2)	-0.286341 (0.24052) [-1.19053]	-0.073100 (0.12093) [-0.60450]
GDP(-3)	0.116330 (0.24726) [ 0.47049]	-0.066033 (0.12432) [-0.53113]

Рис. 44. Оцінка коефіцієнтів VAR-моделі

Vector Autoregression Estimates		
R-squared	0.982168	0.951609
Adj. R-squared	0.972728	0.925991
Sum sq. resids	2.77E+08	70053154
S.E. equation	4037.526	2029.969
F-statistic	104.0403	37.14534
Log likelihood	-256.2574	-237.6919
Akaike AIC	19.72277	18.34755
Schwarz SC	20.20271	18.82749
Mean dependent	59939.44	10495.56
S.D. dependent	24448.81	7461.854
<hr/>		
Determinant Residual Covariance	6.56E+13	
Log Likelihood (d.f. adjusted)	-506.1282	
Akaike Information Criteria	38.97246	
Schwarz Criteria	39.93234	

Рис. 45. Оцінка якості побудованої VAR-моделі.

Наступним кроком після оцінки системи є її діагностика. У таблиці представлені основні команди для проведення такої діагностики.

Меню	Опис
<b>View→Lag Structure→AR Roots Table / View→Lag Structure→AR Roots Graph</b>	Виводить корені поліному для перевірки стаціонарності моделі у табличній або графічній формі. Для стаціонарної моделі всі корені мають бути за абсолютною величиною менші за 1.
<b>View→Lag Structure→Pairwise Granger Causality Tests</b>	Проводить тест Гренджера, за яким визначається, чи може ендогенна змінна вважатися екзогенною. Для кожного рівняння показується статистика Вальда для сукупної значимості кожної з інших лагових ендогенних змінних у заданому рівнянні. Статистика в останньому рядку дозволяє перевірити гіпотезу про сукупну значимість всіх лагових ендогенних змінних у рівнянні.
<b>View→Lag Structure→Lag Exclusion Tests</b>	Проводить тестування щодо доцільності виключення лагів однієї чи всіх змінних.



<b>View→Lag Structure→Lag Length Criteria</b>	Шляхом порівняння декількох критеріїв дозволяє обрати оптимальну кількість лагів в системі.
<b>View→Residual Tests→Correlograms</b>	Повертає попарні кореляційні значення для всіх лагів змінних системи.
<b>View→Residual Tests→Portmanteau Autocorrelation Test</b>	Проводить тест гіпотези про відсутність автокореляції залишків.
<b>View→Residual Tests→Autocorrelation LM Test</b>	Проводить тест гіпотези про відсутність автокореляції залишків.
<b>View→Residual Tests→Normality Test</b>	Проводить тест гіпотези про нормальний розподіл залишків.
<b>View→Residual Tests→White Heteroskedasticity Test</b>	Проводить тест Уайта гіпотези про наявність гетероскедастичності залишків.
<b>View→Impulse Response...</b>	Виводить інформацію про вплив шоку однієї зі змінних на інші ендогенні змінні. За допомогою команди можна обрати форму необхідної інформації (графік, таблиця), вказати список змінних для аналізу, структуру шоку.
<b>View→Variance Decomposition...</b>	Виводить інформацію про відносну важливість кожного з шоків при впливі за змінні VAR-моделі.

Більшість макроекономічних даних мають одиничні корені, неврахування чого може призвести до неправильної оцінки моделі. Метою **тесту Йохансена** на коінтеграцію є перевірка гіпотези про коінтегрованість нестационарних часових рядів. Для проведення тестування користувач має обрати команду **View→Cointegration Test...**, у вікні вказати припущення про форму тренду змінних (рис. 46), та, при необхідності, обмеження моделі.

Якщо користувач впевнений у коінтегрованості нестационарних процесів, то він може побудувати *VEC*–модель. Основна відмінність від *VAR*–моделі полягає у вказуванні користувачем форми коінтеграції та необхідних обмежень для моделі. Основні параметри залишаються такими ж, як і для *VAR*–моделі, проте у закладінці *Cointegration* вказується форма коінтеграції, у закладінці – *VEC Restrictions* – необхідні обмеження на коефіцієнти моделі.

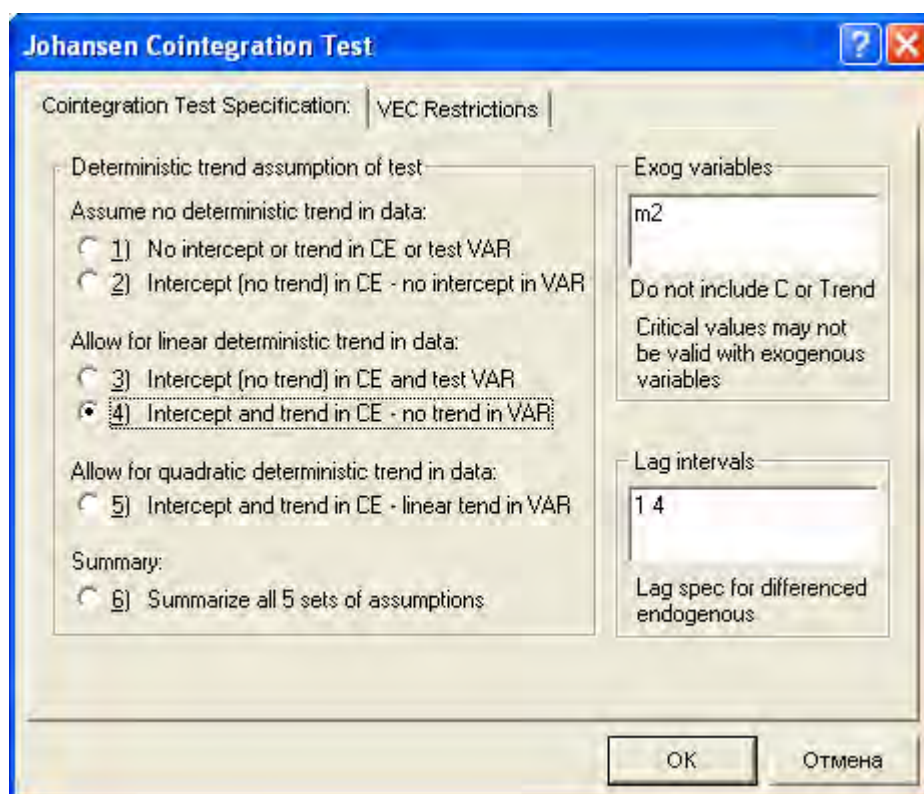


Рис. 46. Проведення тестування на наявність коінтеграції.

### Інші методи оцінювання

Система EViews пропонує й інші види методів оцінки моделей в залежності від характеристик вхідних змінних. Перелік методів та опис моделей, для яких вони застосовуються, наведено у таблиці.

Метод	Опис методу
TSLS	Двоетапний метод найменших квадратів, використовується, наприклад, у випадку наявності у правій частині рівняння ендогенних змінних.
GMM	Узагальнений метод моментів, використовується як альтернатива методу максимальної правдоподібності у випадку невідомої функції розподілу збурень.
Binary	Використовується для оцінки моделей з бінарними залежними змінними.
Ordered	Використовується для оцінки моделей, в яких залежні змінні приймають лише певні значення.
Censored	Використовується для оцінки моделей, в яких пропущена чи невідома певна частина інформації.
Count	Використовується для оцінки моделей, в яких всі змінні приймають лише цілі числа.

## Лабораторна робота

**Задача 7.1.** Для часових рядів з бази даних провести тестування на наявність попарної коінтеграції за методами:

- Інґа-Гренджера;
- Йохансена.

**Задача 7.2.** Побудувати VAR–моделі з двома, трьома та чотирма змінними. На основі різноманітних критеріїв вибрати оптимальну довжину лагів. Отримати прогнози, оцінити похибку прогнозування.

**Задача 7.3.** Провести тестування на необхідність включення ще однієї змінної за допомогою критерію Гренджера.

**Задача 7.4.** Для кожної VAR–моделі (див. зад. 7.2) побудувати імпульсно-відповідні функції.

**Задача 7.5.** Побудувати VAR–моделі (див. зад. 7.2), використовуючи замість змінних їх перші різниці. Визначити адекватність отриманих моделей.

**Задача 7.6.** Для кожної з моделей (див. зад. 7.2, 7.5) побудувати прогнози на наступні періоди, визначити точність прогнозування.

**Задача 7.7.** За допомогою аналізу змінних та проведення імпульсно-відповідного аналізу зробити висновок про доцільність та наслідки впливу на інші макроекономічні показники:

- монетарних шоків (збільшення грошової маси в економіці країни);
- інвестиційних шоків (збільшення державних витрат на капітальне будівництво);
- підвищення ставки рефінансування НБУ;
- збільшення вкладів населення у банках;
- зростання експорту товарів за рахунок сприятливої кон'юнктури на міжнародних ринках;
- адміністративну зміну курсу національної валюти;
- збільшення вартості ціни імпортованих товарів.

**Задача 7.8.** На основі існуючої бази даних макроекономічних показників підібрати та оцінити найкращу систему одночасних рівнянь, що має враховувати рівняння–тотожності (наприклад, ВВП дорівнює сумі державного споживання, інвестицій, державних витрат та чистого експорту) та рівняння поведінки (наприклад, залежність споживання від рівня ВВП поточного та минулих періодів).

**Задача 7.9.** На основі існуючої бази даних макроекономічних показників оцінити систему виду:

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \varepsilon_{1t},$$

$$I_t = \beta_0 + \beta_1 R_t + \beta_2 I_{t-1} + \varepsilon_{2t},$$

$$R_t = \gamma_0 + \gamma_1 Y_t + \gamma_2 M_t + \varepsilon_{3t},$$

$$Y_t = C_t + I_t + G_t.$$

## РОЗДІЛ 8. СУЧАСНЕ ПРОГНОЗУВАННЯ

### Теоретична частина

#### Комбінування прогнозів

На практиці дослідники застосовують багато методів прогнозування, користуючись своїми уподобаннями, навичками, володінням програмним забезпеченням, замовленням на застосування визначеної методики тощо. Звичайно, при використанні будь-якого методу спеціалісти намагаються добитися мінімальної похибки при прогнозуванні. Іноді буває, що один з методів, який добре зарекомендував себе в минулому, дає погані прогнози і навпаки. Щоб застрахуватися від подібних ситуацій, а також поліпшити точність прогнозування необхідно використовувати комбінації прогнозів.

Найбільш відомими є дві методики :

- 1) дисперсійно-коваріаційний метод, що дозволяє зводити декілька незміщених прогнозів в лінійну комбінацію з найменшою дисперсією, ваги якої залежать від дисперсій та коваріацій похибок прогнозів;
- 2) регресійний метод, який є узагальненням дисперсійно-коваріаційного на випадок зміщеності прогнозів.

#### ДИСПЕРСІЙНО-КОВАРІАЦІЙНИЙ МЕТОД

Нехай існує два незміщених прогнози на період  $t$ :  $F_{1t}$  та  $F_{2t}$ . Нехай також дисперсії прогнозів  $\sigma_1^2$  та  $\sigma_2^2$ , коваріація  $\sigma_{12}$ . Новий незміщений прогноз будуватиметься за правилом:

$$F_{3t} = \lambda F_{1t} + (1 - \lambda) F_{2t}.$$

Дисперсія похибки становитиме:

$$\sigma_3^2 = \lambda^2 (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}) - 2\lambda (\sigma_2^2 - \sigma_{12}) + \sigma_2^2.$$

Мінімізуючи вираз по  $\lambda$ , отримуємо

$$\lambda = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}.$$

Звідси

$$\sigma_3^2 = \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{(\sigma_1 - \rho \sigma_2)^2 + \sigma_2^2 (1 - \rho^2)}, \text{ де } \rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2}.$$

Оскільки  $(\sigma_1^2 - \sigma_3^2) \geq 0$  та  $(\sigma_2^2 - \sigma_3^2) \geq 0$ , то комбінований прогноз є не гіршим, ніж найкращий з двох прогнозів.

На практиці часто значення дисперсій та коваріацій похибок прогнозу є невідомими, тому замість них використовують їх оцінки. Таким чином обираються ваги для побудови нового комбінованого прогнозу.

У випадку  $N$  прогнозів,  $N$ -мірний вектор оптимальних вагів визначається за формулою:

$$\Lambda = \frac{V^{-1}I}{I'V^{-1}I},$$

де  $V$  – коваріаційна матриця похибок розмірності  $N \times N$ ,  
 $I$  –  $N$ -мірний вектор одиниць.

З вищесказаного робимо висновок, використання дисперсійно-коваріаційної комбінації є кращою, ніж вибір найкращого з прогнозів з найменшою дисперсією.

### РЕГРЕСІЙНИЙ МЕТОД

Регресійний метод є узагальненням попереднього методу, який інтерпретується як оцінка коефіцієнтів регресійного рівняння виду:

$$F_t = \beta_0 + \beta_1 F_{1t} + \beta_2 F_{2t} + \dots + \beta_N F_{Nt} + \varepsilon_t.$$

Новий комбінований прогноз  $F_t$  є лінійною комбінацією  $N$  прогнозів. Коефіцієнти  $\beta_i$ ,  $i = 0, 2, \dots, N$  оцінюються за методом найменших квадратів. Якщо всі прогнози є незміщеними, то доданок  $\beta_0$  можна опустити. В цьому випадку оцінки коефіцієнтів будуть співпадати з оцінками вектора  $\Lambda$  з попереднього методу.

### Проблема дезагрегування часових рядів

Більшість макроекономічної інформації подається щоквартально або щорічно. Тому іноді виникає проблема співставлення таких даних з іншими, що мають, наприклад, місячну структуру. Можлива також і така ситуація: для одного чи декількох років не подана щоквартальна розбивка. Таким чином, для того щоб можна було використовувати всю наявну інформацію, необхідно вміти розбивати, наприклад, щорічні дані на квартальні.

Припустимо, що ми розглядаємо часовий ряд  $\{Y_T\}$  з річною структурою даних. Наша задача – утворити новий часовий ряд  $\{X_{4T}\}$ , який буде відповідати за значення процесу у кварталах кожного року. Очевидно, що така розбивка часового ряду повинна проводитися за умови

$$y_t = x_{4t-3} + x_{4t-2} + x_{4t-1} + x_{4t}.$$

Розглянемо основні методи, які застосовуються для утворення часового ряду  $\{X_{4T}\}$ .

### ПРОЦЕНТНЕ ВІДНОШЕННЯ

Якщо відома розбивка якогось року по кварталах, то ми можемо використати її для утворення нового ряду з квартальною структурою. Нехай для деякого року  $y_\tau$  відомі квартальні значення  $d_{\tau 1}$ ,  $d_{\tau 2}$ ,  $d_{\tau 3}$ ,  $d_{\tau 4}$ . Тоді новий ряд будується за правилом:

$$x_{4t-4+i} = y_\tau \cdot \alpha_i, \quad i = \overline{1, 4}, \quad \text{де } \alpha_i = \frac{d_{\tau i}}{y_\tau}.$$

Цей метод утворить новий ряд, який матиме пропорційну структуру. Якщо дослідник має квартальну розбивку по декількох роках, номери яких складають множину  $S = \{t_1, t_2, \dots, t_n\}$ , то

$$\alpha_i = \sqrt[n]{\prod_{\tau \in S} \frac{d_{\tau i}}{y_{\tau}}}, i = \overline{1, 4}.$$

Якщо множина  $S$  є пустою, тобто невідомо розбивки по кварталах жодного року, то використовується елементарне усереднення:

$$\alpha_i = \frac{1}{4}, i = \overline{1, 4},$$

яке виключає сезонні коливання.

### ПОЛІНОМІАЛЬНА ІНТЕРПОЛЯЦІЯ

Цей метод спочатку утворює новий ряд за правилом

$$z_j = \sum_{t=1}^j y_t, j = \overline{1, T}.$$

Ряд  $\{Z_T\}$  представляє собою акумулятивні суми початкового ряду. Після цього, послідовно через чотири точки ряду  $\{Z_T\}$  будується кубічна інтерполяція, тобто знаходяться коефіцієнти функції  $f(t) = at^3 + bt^2 + ct + d$ . Наприклад, для розбиття  $y_3$  нам необхідно побудувати інтерполяцію по точках  $z_1, z_2, z_3, z_4$ :

$$\begin{cases} f(1) = z_1, \\ f(2) = z_2, \\ f(3) = z_3, \\ f(4) = z_4. \end{cases}$$

Тепер підраховуємо значення  $f(2), f(2.25), f(2.5), f(2.75), f(3)$ .

Тоді

$$\begin{aligned} x_9 &= f(2.25) - f(2), \\ x_{10} &= f(2.5) - f(2.25) \\ x_{11} &= f(2.75) - f(2.5) \\ x_{12} &= f(3) - f(2.75). \end{aligned}$$

Зазначимо, що сума

$$x_9 + x_{10} + x_{11} + x_{12} = f(3) - f(2) = z_3 - z_2 = y_3.$$

### BFL-FD ТА BFL-SD МЕТОДИ

Математично цей метод записується у формі:

$$\sum_{t=2}^{4T} (\Delta x_t)^2 \rightarrow \min$$

за умови

$$\sum_{j=1}^4 x_{4t-4+j} = y_t, t = \overline{1, T}, \text{ де } \Delta x_t = x_t - x_{t-1}.$$

Деякою модифікацією є мінімізація суми квадратів других різниць:

$$\sum_{t=3}^{4T} (\Delta x_t - \Delta x_{t-1})^2 \rightarrow \min$$

за умови

$$\sum_{j=1}^4 x_{4t-4+j} = y_t, t = \overline{1, T}.$$

### Етапи сучасного прогнозування

Сучасне прогнозування неможливе без широкого поєднання статистичних методів та експертних оцінок. В останні роки статистичні методи набули достатньо великого розвитку. Необхідно виділити 5 етапів побудови прогнозів:

1. Постановка задачі.
2. Отримання інформації.
3. Виконання методів прогнозування.
4. Оцінка методів прогнозування.
5. Використання прогнозів.

Перший етап складається з постановки цілей дослідження, визначення взаємовпливу прогнозів та зовнішнього середовища. Наступним кроком є структурування задачі, виділення її компонентів, детермінованих (тренд, сезонні коливання тощо) та випадкових. Останні можуть бути викликані як одночасними подіями, так і цілими ланцюгами зовнішніх впливів.

На другому етапі слід виділити і визначити джерела отримання інформації. При цьому доцільно:

- а) уникати використання недостовірної або помилкової інформації, оскільки навіть її наявність може призвести до невірних висновків;
- б) визначити, яка саме інформація є важливою, причому слід використовувати останню наявну інформацію, оскільки вона може змінювати попередню і не узгоджуватися з нею;
- в) користуватися не одним, а декількома джерелам;
- г) використати досвід експертів.

Наступним кроком є так звана “чистка” даних. Вона вимагає виключення з інформації помилок, зміни значень внаслідок зміни визначень економічних понять, інфляції тощо. Крім того, необхідно виключити минулі систематичні та несистематичні зміни, сезонні коливання, відомі шоки. В деяких випадках такі зміни фіксуються графічним аналізом.

На третьому етапі слід обрати відповідні експертні та статистичні методи прогнозування. Якщо прогнози мають політичний вплив на економіку, то вибір значно ускладнюється. Для вибору найкращого методу залучаються експерти, які б на основі свого досвіду проранжували методи. Краще використовувати



статистичні методи, які для рівних початкових умов видають однакові прогнози, що значно полегшує подальший аналіз.

При застосуванні статистичного прогнозування слід використовувати прості, надійні методи, особливо в умовах високої невпевненості в майбутньому. Не варто також забувати про зміну прогнозного значення у зв'язку з майбутніми передбаченими подіями. Не слід намагатися виділяти циклічні коливання, оскільки на сьогоднішній час ще не розроблено загально прийнятного методу.

При використанні експертних методів слід вимагати від експертів повного розуміння питання і їх письмового прогнозу з його обґрунтуванням. Крім того, лише залучення достатньої кількості експертів призводить до покращання точності прогнозів. Можна вимагати від них також розробляти сценарні підходи, після чого проводити ранжування експертів по точності їх прогнозів.

При одночасному застосуванні експертних та статистичних методів потрібно заздалегідь визначити правила їх комбінування. Комбінування прогнозів є особливо нагальним, коли потрібно отримати не найточніший з прогнозів, а запобігти великій похибці.

На четвертому етапі дослідник повинен обрати один з декількох методів. Найкращий метод потрібно обрати після повного аналізу залишків, порівняння критерії точності, а також витрат на застосування методів. При порівнянні методів іноді використовують надійні інтервали. Доцільним є й обґрунтування дослідником причин, за яких прогноз може бути неточним. Такі дані дозволяють полегшувати майбутній аналіз.

На п'ятому етапі необхідно подати результати дослідження до замовника. Данні повинні бути представлені у зручній для сприйняття формі, обґрунтування методів подано просто і точно, всі припущення, зроблені дослідником, мають бути чітко окреслені. Крім загальних результатів роботи слід подати міркування щодо використання адаптивних методів, коли дослідник має змогу змінювати коефіцієнти моделі в процесі роботи.

### ***Лабораторна робота***

**Задача 8.1.** На основі отриманих прогнозів для Ваших часових рядів (розділи 1–7) побудувати комбінований прогноз на основі:

- методу усереднення;
- дисперсійно-коваріаційного методу;
- регресійного методу.

Для кожного з методів підрахувати похибку прогнозування.

**Задача 8.2.** Проранжувати для кожного з часових рядів всі методи за точністю прогнозування на поточний рік.

**Задача 8.3.** Підготувати заключний звіт про виконання всіх завдань з економічним аналізом відповідних часових рядів.

## ПРОГРАМУВАННЯ У EVIEWS

Будь-яку операцію у EViews можна здійснити за допомогою команди, що записується у верхній частині вікна під головним меню системи. Наприклад, запис

*ls gdp c gdp(-1)*

створює та оцінює регресії  $gdp_t = \beta_0 + \beta_1 gdp_{t-1} + \varepsilon_t$  за допомогою методу найменших квадратів. Кількість команд у такому вікні не обмежена, проте всі вони зберігаються лише до кінця поточного сеансу з системою. Для того, щоб створювати набори команд, які можуть зберігатися як окремі файли, передбачені файли-програми (розширення *.prg*).

Створення нової програми здійснюється аналогічно до створення нової робочої книги (**File**→**New**→**Program**). При цьому відкривається вікно, в якому користувач має можливість записувати текст програми. Для зручності роботи вікно містить кнопки для редагування, пошуку, заміни тексту. Запис програми на диск здійснюється за допомогою кнопок **Save** та **SaveAs**. Кнопка **Run** запускає програму на виконання. Зупинити виконання команди можна в будь-який момент клавішею *F1*.

Для написання програм потрібно розглянути основні характеристики мови програмування EViews.

### Змінні

Змінні у програмах зустрічаються двох типів: числові та строкові. Для оголошення числової змінної перед її ім'ям ставиться знак **!**, а перед строковою – **%**. Для надання змінній певного значення воно вказується після знаку **=**. Наступні приклади ілюструють створення двох змінних:

*!a=7*

*%s="Текстова змінна"*

Для оголошення чи створення об'єктів слід вказати тип об'єкта та його назву.

Загальним записом для створення об'єктів є запис

*object\_type object\_name*

Типом об'єкту може бути: *coef, equation, graph, group, matrix, model, pool, rowvector, sample, scalar, series, sspace, sym, system, table, text, var, vector*. Створений об'єкт бути розташований у активній робочій книзі. Приклади створення об'єктів наведені у таблиці.

Запис	Коментар
<i>series lgdp</i>	створює новий часовий ряд, який називається <i>lgdp</i> .
<i>equation eql</i>	створює нове економетричне рівняння з ім'ям <i>eql</i>
<i>matrix(5,5) x</i>	створює матрицю <i>x</i> з 5 рядками та 5 стовпчиками

Для звернення до об'єктів використовується синтаксис  
*action(act\_opt) object.view\_proc(view\_proc\_opt) arg\_list*

де

<i>action</i>	одне з ключових слів, що визначають необхідну дію ( <i>do</i> , <i>freeze</i> , <i>print</i> , <i>show</i> )
<i>act_opt</i>	необов'язковий параметр, що визначає виконання дії
<i>object</i>	ім'я об'єкта, над яким потрібно провести дію
<i>view_proc</i>	необхідна процедура об'єкта
<i>view_proc_opt</i>	необов'язковий параметр, що визначає поведінку об'єкта під час виконання процедури
<i>arg_list</i>	список параметрів, розділених пробілами.

Існує 4 можливих дії:

<i>do</i>	виконує процедуру без відкриття вікна. Якщо вікно об'єкта не відкрите, то процедура не виконується.
<i>freeze</i>	створює таблицю або графік по даних об'єкта.
<i>print</i>	роздруковує вікно об'єкта.
<i>show</i>	показує вікно об'єкта.

Якщо дія не вказана, то виконується дія *show*. Наприклад, запис

*gdp.line*

еквівалентний запису

*show gdp.line*

та призводить до побудови лінії графіку по змінній *gdp*.

Присвоєння об'єктів аналогічно присвоєнню змінних:

*object\_name = expression*

Наприклад, для створення нового часового ряду, який міститиме натуральні логарифми ряду *gdp*, достатньо записати:

*series newgdp*  
*newgdp=log(gdp)*

Якби змінна *newgdp* вже існувала, то перший рядок був би необов'язковим.

EViews дозволяє створювати об'єкти, одночасно присвоюючи їм значення. Наприклад,

*scalar elas = 2*  
*series tr58 = @trend(1958)*  
*group nipa gdp cons inv g x*  
*equation cnsfnc2.ls log(cons)=c(1)+c(2)\*yd*  
*vector beta = @inverse(x\*x)\*(x\*y)*

В той же час можна створювати об'єкти без імені, наприклад,

*ls y c x1 x2*

Такий запис створює нове економетричне рівняння, оцінює його, проте не вставляє об'єкт до робочої книги.

### Робота з робочими книгами та базами даних

При створенні робочої книги необхідно вказати періодичність, початок та кінець діапазону даних.

Наприклад, запис

*workfile macro1 q 1965:1 1995:4*

створює нову робочу книгу з ім'ям *macro1*, що містить квартальні дані від першого кварталу 1965 року до 4 кварталу 1995 року.

Періодичність вказується однією англійською літерою: *a* – річні, *s* – на піврічні, *q* – квартальні, *m* – місячні, *w* – тижневі, *d* – добові (5 або 7 днів на тиждень), *u* – неперіодичні.

З робочою книгою можна робити операція збереження на диску (команда *save*), закриття (*close*), завантаження нової книги (*open*).

Робота з базами даних аналогічна роботі з робочою книгою. Найбільш важливі команди роботи з базою даних представлені у таблиці.

Команда	Опис	Приклад
<i>dbcreate</i>	створює нову базу даних	<i>dbcreate mydata1</i>
<i>db</i>	відкриває базу даних за заданою адресою	<i>db c:\evdata\usdb</i>
<i>dbopen</i>	відкриває базу даних у поточній папці	<i>dbopen findat</i>
<i>dbrename</i>	змінює ім'я бази даних зі старого на нове	<i>dbrename temp1 newmacro</i>
<i>dbdelete</i>	знищує базу даних за заданою адресою	<i>dbdelete c:\data\usmacro</i>
<i>dbcopy</i>	копіює базу даних	<i>dbcopy c:\evdata\macro1 a:\macro1</i>

Головні операції з об'єктами бази даних представлені у таблиці

Команда	Опис
<i>copy</i>	створює копію об'єкта чи декількох об'єктів
<i>rename</i>	змінює ім'я об'єкта
<i>delete</i>	знищує об'єкт
<i>store</i>	зберігає об'єкт у робочій книзі або базі даних
<i>fetch</i>	завантажує об'єкт з бази даних

## Оператори та цикли

Часто виникає потреба у непослідовному виконанні команд програми. У таких випадках існує необхідність у використанні циклів.

**Умовний оператор** дозволяє перевірити виконання однієї або декількох умов. Починається він зі службового слова *If*, за яким слідує умова. Якщо умова справджується, то далі виконуються команди, що стоять за словом *Then*, у протилежному випадку – за словом *Else*. Закінчується умовний оператор словом *End If*. Особливо зручною є можливість створення вкладених умовних операторів.

Наведений приклад перевіряє середнє значення змінної *gdp*. Якщо воно перевищує 50 000, створюється нова змінна, що містить елементи *gdp*, зменшені на знайдене середнє значення, у протилежному випадку, нова змінна *newgdp* міститиме елементи *gdp*, зменшені на 50 000.

```
if @mean(gdp)>50000 then
series newgdp = gdp-@mean(gdp)
else
series newgdp = gdp-50000
endif
```

Для багатократного виконання одного чи декількох операторів призначений **цикл безумовного проходу** *For ... Next*. Наприклад, запис

```
series my
for !j=1 to 10
my(!j)=100
net
```

створює змінну *my*, та заповнює перші 10 значень числом 100. При потребі цикл може бути вкладеним.

Якщо не вказуються додаткові параметри, то ітеративна змінна збільшується при виконанні кожного циклу на 1. При необхідності змінювати крок на іншу величину, застосовується службове слово *step*, наприклад,

```
for !j=0 to 10 step 3
series z = z/!j
net
```

Цикл також може бути записаний за допомогою рядкових змінних. У такому випадку перебираються всі змінні, для яких виконується відповідна команда. Наприклад, програма

```
for %y gdp g inv
equation {%y}_eq.ls %y c {%y}(-1) @trend
net
```

створює та оцінює 3 економетричних рівняння. Перша регресія виду  $gdp_t = \beta_0 + \beta_1 gdp_{t-1} + \beta_2 t + \varepsilon_t$  називається *gdp\_eq*, друга регресія виду  $g_t = \beta_0 + \beta_1 g_{t-1} + \beta_2 t + \varepsilon_t$  називається *g\_eq*, нарешті, третя регресія виду  $inv_t = \beta_0 + \beta_1 inv_{t-1} + \beta_2 t + \varepsilon_t$  має ім'я *inv\_eq*.

Існує можливість вказувати навіть окремі вибірки для кожного рівняння. Запис

```
for %1 %2 %3 1955:1 1960:4 early 1970:2 1980:3 mid 1975:4 1995:1 late
  smpl %1 %2
  equation {%3}eq.ls gdp c inv
next
```

дозволяє оцінити три регресії (*earlyeq*, *mideq*, *lateeq*) одного виду ( $gdp_t = \beta_0 + \beta_1 inv_t + \varepsilon_t$ ) на 3 різних вибірках.

**УМОВНИЙ ЦИКЛ** *While...Wend* виконується доти, поки залишається вірною умова у пропозиції *While*. Наприклад, програма

```
!i = 1
while !i <= 25
  scalar a{!i} = !i
  !i = !i + 11
wend
```

створює 3 скаляри з іменами *A1*, *A12*, *A23*.

Для негайного виходу з циклів *For...Next* та *While...Wend* використовується службове слово *exitloop*, при виконанні якого управління переходить на наступну команду за словом *Next* або *Wend*.

## Процедури

Процедури являють собою набір команд, що можуть виконуватися багато разів з різними параметрами. Вони дозволяють зробити текст програми більш легким для сприйняття, оптимізувати виконання необхідних дій.

Процедура починається зі службового слова *subroutine*, за яким слідує назва процедури. При необхідності в дужках описуються параметри підпрограми. Завершується процедура словом *endsub*. Наприклад, процедура *power* підносить кожен елемент вектора *u* до степеня *p* і записує результат у вектор *v*:

```
subroutine power(series v, series u, scalar p)
  v = u^p
endsub
```

Для виклику процедури у тексті програми слід записати:

```
call power(gdp2,gdp,2)
```

При використанні процедур слід розрізняти локальні та глобальні змінні. Як правило, всі створені об'єкти та змінні є глобальними, тобто доступними з

будь-якої процедури. Якщо назва параметра процедури співпадає з назвою глобальної змінної, то в межах процедури змінна вважається локальною.

Іноді потрібно у процедурі тимчасово створювати велику кількість змінних та обробляти їх. В таких випадках використовують локальні процедури, які автоматично знищують всю створену у процедурі інформацію по завершенні їх роботи. Для оголошення процедури такого типу слід вказувати службове слово *local* після оголошення процедури, наприклад,

```
subroutine local power
```

```
...
```

```
endsub
```

Перервати виконання процедури та повернутися до процедури вищого рівня чи головної програми можна за допомогою команди *return*.

### Функції

Робота з EViews значно спрощується через наявність великої кількості функцій, що виконують стандартні операції. Їх можна використовувати як при написанні програм, так і при генерації нових змінних при ручному керуванні системою. Всі функції можна розбити на декілька категорій.

**Математичні функції** здійснюють стандартні операції по знаходженню математичних та тригонометричних функцій.

Назва	Опис
@abs(x), abs(x)	знаходить модуль числа
@acos(x)	знаходить арккосинус
@asin(x)	знаходить арксинус
@atan(x)	знаходить арктангенс
@ceiling(x)	округлює число до найближчого більшого цілого числа
@cos(x)	знаходить косинус
@exp(x), exp(x)	підносить число <i>e</i> до степеня <i>x</i>
@fact(x)	знаходить факторіалу числа
@factlog(x)	знаходить логарифм натуральний факторіалу числа
@floor(x)	округлює число до найближчого меншого цілого числа
@inv(x)	знаходить обернене число $1/x$
@log(x), log(x)	обчислює натуральний логарифм числа <i>x</i>
@mod(x, y)	знаходить залишок від ділення <i>x</i> на <i>y</i>
@nan(x, y)	повертає <i>x</i> , якщо <i>x</i> не є порожнім, інакше повертає <i>y</i>
@recode(s, x, y)	якщо умова <i>s</i> виконується, то повертається <i>x</i> , інакше – <i>y</i>
@round(x)	округлює до найближчого цілого
@sin(x)	знаходить синус
@sqrt(x), sqr(x)	знаходить квадратний корінь
@tan(x)	знаходить тангенс



**Текстові функції** дозволяють здійснювати операції з строковими змінними.

Функція	Опис
@val	перетворює текст у число
@str	перетворює число у текст
@left	повертає ліві символи тексту
@mid	повертає символи з середини тексту
@right	повертає праві символи тексту
@otod	повертає дату у форматі рядка, яка є зсувом від початку активного діапазону
@dtoo	повертає кількість періодів від початку активного діапазону

**Функції часових рядів** перетворюють значення часового ряду.

Назва	Опис
d(x)	знаходить різницю першого порядку
d(x,n)	знаходить різницю n-го порядку
d(x,n,s)	знаходить різницю n-го порядку з врахуванням сезонного компонента
dlog(x)	знаходить різницю першого порядку у логарифмах значень
dlog(x,n)	знаходить різницю n-го порядку у логарифмах значень
dlog(x,n,s)	знаходить різницю n-го порядку у логарифмах значень з врахуванням сезонного компонента
@movav(x,n)	застосовує метод усереднення з кроком n
@movsum(x,n)	повертає суми значень ряду за n кроків
@pch(x)	повертає відсоткову зміну за один період
@pcha(x)	повертає відсоткову зміну за один період, враховуючи структуру даних
@pchy(x)	повертає відсоткову зміну за один рік
@seas(n)	створює фіктивні змінні
@trend, @trend(n)	створює трендові змінні

**Функції дат** дозволяють визначати інформацію про час отримання значень часового ряду.

Назва	Опис
@day	повертає день спостереження для добових чи тижневих даних
@elem(x,"d"), @elem(x,s)	повертає значення часового ряду у відповідний період
@month	повертає номер місяця для відповідного значення ряду
@quarter	повертає номер кварталу для відповідного значення ряду
@year	повертає номер року для відповідного значення ряду

**Функції описової статистики** дозволяють отримати основні числові характеристики часового ряду.

Назва	Опис
@cor(x,y[,s])	повертає коефіцієнт кореляції
@cov(x,y[,s])	повертає коефіцієнт коваріації
@inner(x,y[,s])	повертає скалярний добуток векторів
@obs(x[,s])	повертає кількість спостережень
@mean(x[,s])	повертає вибіркове середнє
@median(x[,s])	повертає медіану вибірки
@min(x[,s])	повертає мінімальне значення ряду
@max(x[,s])	повертає максимальне значення ряду
@quantile(x,q[,m,s])	повертає квантиль вибірки
@stdev(x[,s])	повертає середньоквадратичне відхилення
@sum(x[,s])	повертає суму ряду
@sumsq(x[,s])	повертає суму квадратів ряду
@var(x[,s])	повертає дисперсію вибірки

**Функції розподілів** дозволяють користувачу отримувати значення відомих статистичних розподілів, не використовуючи статистичні таблиці. Для будь-якого розподілу у EViews передбачено 4 функції. Перша повертає значення функції розподілу (починається з @c), друга повертає значення функції щільності (@d), третя повертає обернену функцію розподілу (@q), нарешті, четверта повертає випадкове число, згенероване за заданим розподілом (@r).

Розподіл	Функції
$\beta$ -розподіл	@cbeta(x,a,b), @dbeta(x,a,b), @qbeta(p,a,b), @rbeta(a,b)
Біноміальний розподіл	@cbinom(x,n,p), @dbinom(x,n,p), @qbinom(s,n,p), @rbinom(n,p)
$\chi^2$ —розподіл	@cchisq(x,v), @dchisq(x,v), @qchisq(p,v), @rchisq(v)
Експоненціальний розподіл	@cexp(x,m), @dexp(x,m), @qexp(p,m), @rexp(m)
$F$ –розподіл Фішера	@cfdist(x,v1,v2), @dfdist(x,v1,v2), @qfdist(p,v1,v2), @rfdist(v1,v1)
$\Gamma$ –розподіл	@cgamma(x,b,r), @dgamma(x,b,r), @qgamma(p,b,r), @rgamma(b,r)
Розподіл Лапласа	@claplace(x), @dlaplace(x), @qlaplace(x), @rlaplace
Логістичний розподіл	@clogistic(x), @dlogistic(x), @qlogistic(p), @rlogistic
Логнормальний розподіл	@clognorm(x,m,s), @dlognorm(x,m,s), @qlognorm(p,m,s), @rlognorm(m,s)

Розподіл	Функції
Обернений біноміальний розподіл	@cnegbin(x,n,p), @dnegbin(x,n,p), @qnegbin(s,n,p), @rnebin(n,p)
Нормальний розподіл	@cnorm(x), @dnorm(x), @qnorm(p), @rnorm, nrnd
Розподіл Пуассона	@cpoisson(x,m), @dpoisson(x,m), @qpoisson(p,m), @rpoisson(m)
Розподіл Парето	@cpareto(x,a,k), @dpareto(x,a,k), @qpareto(p,a,k), @rpareto(a,k)
<i>t</i> –розподіл Стюдента	@ctdist(x,v), @dtdist(x,v), @qtdist(p,v), @rtdist(v)
Рівномірний розподіл	@cunif(x,a,b), @dunif(x,a,b), @qunif(p,a,b), @runif(a,b), rnd
Розподіл Вейбула	@cweib(x,m,a), @dweib(x,m,a), @qweib(p,m,a), @rweib(m,a)

### ІНШІ МОЖЛИВОСТІ

Часто при програмуванні доводиться змінювати поточну вибірку за допомогою команди *simpl*, наприклад, наступним чином:

```
simpl 1968:5 1992:12
```

Іноді програма має помилки, які варто обробляти в автоматичному режимі. Для отримання коду помилки існує функція @errorcount. Отриманий код дозволяє автоматично обробляти помилки програми.

Зупинити виконання програми у будь-який момент можна за допомогою команди *stop*, вийти з системи EViews – за допомогою команди *exit*.

Часто доводиться виконувати програми, які вже записані в інших файлах. Тоді їх потрібно додати до основної програми за допомогою інструкції *include*:

```
include c:\programs\powers.prg
```

## ЗАВДАННЯ ДЛЯ САМОСТІЙНОЇ РОБОТИ З КУРСУ

**Завдання 1.** Написати звіт про роботу та основні досягнення відомого економіста, що має значні здобутки в теорії та практиці аналізу часових рядів. Слід зазначити його коротку біографію, основні праці, результати, їхнє застосування в сучасній економіці, здобуті нагороди (3-5 сторінок).

**Завдання 2.** Розібратися та підготувати інформацію про один з відомих методів аналізу часових рядів.

**Завдання 3<sup>1</sup>.** З обраних на початку семестру 5 часових рядів вибрати 1, для якого реалізувати **ВСІ** вивчені протягом семестру методи прогнозування. Результатом роботи має бути:

- дані часового ряду (назва, періодичність, безпосередньо дані);
- зведена таблиця прогнозних значень за ВСІМА методами на 4-6 періодів;
- зведена таблиця помилок прогнозування (на основі даних попереднього року);
- обґрунтований вибір найкращої моделі прогнозування;
- повна інформація по обраній моделі, її розрахунку, реалізації, графічному представленню тощо.

### Варіанти завдання 1.

1. Фелпс (Phelps E.S.)
2. Прескотт (Prescott E.C.)
3. Лукас (Lucas R.E.)
4. Дж. Тейлор (Taylor J.)
5. Фулер (Fuller W.A.)
6. Сімс (Sims C.A.)
7. Грейнджер (Granger C.W.J.)
8. Уотсон (Watson M.W.)
9. Хансен (Hansen L.P)
10. Діболд (Diebold F.X.)
11. Юл (Yule G.U.)
12. Сток (Stock J.H.)
13. Нельсон (Nelson C.R.)
14. Харві (Harvey A.C.)
15. Хендрі (Hendry D.F.)
16. Гамільтон (Hamilton J.D.)
17. Болерслев (Bollerslev T.)
18. Чоу (Chow G.C.)

---

<sup>1</sup> Завдання 3 виконується тільки для стаціонарної форми навчання.

19. Енгл (Engle R.F.)
20. Хілдрет (Hildreth C.)
21. Йохансен (Johansen S.)
22. Тобін (Tobin)
23. Уайт (White H.)
24. Грін (Greene W.)
25. Рао (Rao P.)

## **Варіанти завдання 2.**

1. Знаходження максимуму функції правдоподібності з обмеженнями.
2. IVAR моделі: системи часових рядів з цілими значеннями.
3. Імпульсний аналіз VAR – моделей.
4. Методи підрахунку коваріаційної матриці та її оберненої для стаціонарних AR процесів.
5. Фільтр Кальмана.
6. Метод максимальної правдоподібності при знаходженні невідомих коефіцієнтів ARIMA моделей.
7. Основи спектрального аналізу часових рядів.
8. Гармонійні коливання та їх застосування.
9. Проблеми поєднання прогнозів декількох методів з експертними оцінками.
10. Моделі зі змінними трендами.
11. ME – метод.
12. Інструментальні змінні
13. Аналітичні види тренду у макроекономічних процесах
14. DSGE-моделі
15. Триетапний метод найменших квадратів
16. Модель пробіт (probit)
17. Модель логіт (logit)
18. Модель Тобіт
19. Моделі з панельними даними
20. Експертні методи прогнозування
21. Структурні макроекономічні моделі
22. Моделі множинного вибору
23. Тест Хаусмана
24. Методи Census X11 та Census X12
25. Нелінійний метод найменших квадратів

## КОНТРОЛЬНІ РОБОТИ З КУРСУ

### *Контрольна робота для стаціонарної форми навчання*

#### **Завдання:**

За допомогою сайту [www.bank.gov.ua](http://www.bank.gov.ua) створити файл, що містить помісячну інформацію про заданий у варіанті часовий ряд з 2001 по 2005 роки.

1. Перевірити ряд на стаціонарність за допомогою „методу поворотних значень”.
2. Побудувати та оцінити найкращу модель з трендом.
3. Здійснити згладжування за 6 різними методами.
4. Визначити наявність сезонних коливань.
5. Здійснити прогнозування за моделями, що були побудовані в пп. 2-5.
6. Перевести ряд до стаціонарного методом „різниць”. Над отриманим рядом виконати завдання пп.2-6.
7. Обрати і оцінити найкращу *ARIMA*–модель.
8. Обрати і оцінити найкращу *GARCH*–модель.
9. Використовуючи всі наявні дані, побудувати найкращу економетричну модель.
10. Виписати прогнози на 2006 рік за всіма побудованими моделями.

#### **Варіант 1**

Часовий ряд: M0

#### **Варіант 2**

Часовий ряд: M2

#### **Варіант 3**

Часовий ряд: M3

#### **Варіант 4**

Часовий ряд: M1

#### **Варіант 5**

Часовий ряд: Витрати та заощадження населення

#### **Варіант 6**

Часовий ряд: Інвестиції в основний капітал

#### **Варіант 7**

Часовий ряд: Обсяг промислової продукції

**Варіант 8**

Часовий ряд: Обсяг продукції сільського господарства

**Варіант 9**

Часовий ряд: Середньомісячна номінальна заробітна плата працівників

**Варіант 10**

Часовий ряд: Кредити, надані у національній валюті

**Варіант 11**

Часовий ряд: Середній курс гривні до австралійського долара

**Варіант 12**

Часовий ряд: Середній курс гривні до канадського долара

**Варіант 13**

Часовий ряд: Середній курс гривні до євро

**Варіант 14**

Часовий ряд: Середній курс гривні до російського рубля

**Варіант 15**

Часовий ряд: Кредити, надані у іноземній валюті

**Контрольна робота  
для заочної та очно-заочної форм навчання**

**Задача 1.** Провести графічний аналіз кількості скарг на програмне забезпечення фірми (табл.1). За допомогою методу усереднення з кроком 3 визначити зміну кількості скарг. Застосувати до даних метод експоненціального згладжування та подвійного експоненціального згладжування для отримання прогнозу кількості скарг на наступний період. Оцінити надійність прогнозу.

Таблиця 1. Динаміка кількості скарг.

Місяць	Кількість скарг
1	34
2	39
3	38
4	36
5	33
6	32
7	30
8	37
9	37
10	34
11	35
12	33

Місяць	Кількість скарг
13	34
14	33
15	31
16	33
17	32
18	34
19	34
20	38
21	33
22	36
23	38
24	37



**Задача 2.** Провести графічний аналіз доходів підприємства (табл. 2) За допомогою методу усереднення з кроком 4 визначити наявність трендового компонента. Застосувати для прогнозування методи експоненціального та подвійного експоненціального згладжування. Порівняти отримані прогнози на наступний період за двома методами.

Таблиця 2. Динаміка доходів підприємства.

Період	Доходи, тис. грн.	Період	Доходи, тис. грн.
1	512	13	515
2	511	14	514
3	540	15	530
4	545	16	580
5	550	17	489
6	535	18	546
7	560	19	575
8	555	20	560
9	538	21	553
10	541	22	529
11	540	23	571
12	520	24	526

**Задача 3.** Здійснити прогнозування обсягу операцій для двох банків за даними табл. 3. Визначити найбільш привабливий у фінансовому відношенні банк. Визначити ступінь залежності між розмірами банківських вкладів. Знайти оцінку трендового компонента за методом лінійної регресії.

Таблиця 3. Динаміка обсягу операцій, млн. грн.

Період	Банк А	Банк В	Період	Банк А	Банк В
1	558,359	287,806	13	709,568	510,112
2	554,779	260,146	14	687,238	481,750
3	591,981	228,003	15	701,528	427,398
4	571,263	274,378	16	770,424	518,218
5	595,686	249,499	17	751,412	542,211
6	609,378	266,124	18	748,550	496,274
7	660,442	321,598	19	757,496	514,465
8	651,997	364,874	20	774,630	593,817
9	652,856	450,373	21	822,892	570,909
10	640,412	343,303	22	847,170	631,084
11	700,798	460,190	23	825,947	641,359
12	674,255	370,828	24	865,033	657,133

**Задача 4.** За даним офіційного сайту Національного банку України ([www.bank.gov.ua](http://www.bank.gov.ua)) створити часові ряди спостережень за курсом національної валюти відносно канадського долара. Визначити трендовий компонент часового ряду. Застосувати методи експоненціального згладжування та методу Холта-Вінтерса для прогнозування курсу на періоди попереднього року. Підрахувати похибку моделей.

**Задача 5.** За даним офіційного сайту Національного банку України ([www.bank.gov.ua](http://www.bank.gov.ua)) створити часові ряди спостережень за основними статтями платіжного балансу України (експорт товарів та послуг, імпорт товарів та послуг). Побудувати регресію відносно трендового компонента та сезонних змінних. Підрахувати прогнози на наступні періоди.

**Задача 6.** Підготувати звіт по виконаних завданнях.

## ПИТАННЯ НА ЗАЛІК 3 КУРСУ

1. Прогноз. Економічний прогноз.
2. Прийом прогнозування. Метод прогнозування. Методика прогнозування.
3. Прогнозна модель. Система прогнозування. Ефективність прогнозування. Функції економічного прогнозу.
4. Класифікація прогнозів.
5. Послідовність операцій розробки прогнозу.
6. Часові ряди. Порядок аналізу часових рядів. Лаговий оператор.
7. Адитивна та мультиплікативна моделі часових рядів.
8. Міри точності прогнозів.
9. Функція автокореляції. Корелограма. Аналіз корелограми. Часткова автокореляційна функція.
10. Функція правдоподібності. Максимізація функції правдоподібності. Перевірка гіпотез за допомогою функції правдоподібності.
11. Числові характеристики часових рядів. Стаціонарність часових рядів.
12. Метод поворотних точок.
13. Класичні підходи згладжування часових рядів.
14. Методи експоненціального згладжування. Адаптивне згладжування.
15. Несезонна модель Холта–Вінтерса.
16. Адитивна модель із визначенням сезонних коливань.
17. Адитивна модель Холта–Вінтерса.
18. Мультиплікативна модель Холта–Вінтерса.
19. Фільтр Ходріка–Прескотта.
20. Часові ряди з трендом. Функціональні види тренду.
21. Виділення сезонних коливань.
22. Поняття «білого шуму».  $MA(q)$ -процес. Перетворення  $MA$ -процесів.
23. Процес авторегресії.
24.  $ARMA$ -процес.
25. Прогнозування на основі  $ARMA$ -моделей.
26.  $ARIMA$ -процеси. “Випадкове блукання”.
27. Тестування на наявність одиничного кореня.
28. Імпульсний аналіз.
29. Аналіз часових рядів Бокса–Дженкінса.
30. Числові критерії адекватності моделей.
31. Нелінійні процеси авторегресії, порогові моделі, білінійні моделі, гармонічні процеси.
32.  $ARCH$ -процеси та їх оцінка.
33.  $GARCH$ -процеси та їх оцінка.
34.  $TARCH$ -процеси,  $EGARCH$ -процеси,  $ARCH-M$ -процеси та їх оцінка.
35. Моделювання часових рядів при зміні економічної ситуації.
36. Означення  $VAR$ -моделі. Різновиди  $VAR$ -моделей.

37. Оцінка стаціонарних VAR-моделей. Прогнозування на основі VAR-моделей.
38. Структурний аналіз на основі VAR-моделей.
39. VEC-моделі.
40. Імпульсний аналіз VAR-моделей.
41. Причинність за Гренджером.
42. Коінтеграція. Тестування на наявність коінтеграції.
43. Тест Йохансена.
44. Комбінування прогнозів.
45. Нові шляхи прогнозування.
46. Принципи вибору моделей для прогнозування.
47. Експертне прогнозування.
48. Методи вибірових досліджень при прогнозуванні часових рядів.  
Прогнозування на основі нейронних мереж.
49. Проблема дезагрегування часових рядів.
50. Метод поліноміальної інтерполяції.

## ЛІТЕРАТУРА

1. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. – М.: Мир, 1974.
2. Геєць В.М. та ін. Моделі і методи соціально-економічного прогнозування. – Х.: ВД „ІНЖЕК”, 2005. – 396 с.
3. Иванов В.В. Анализ временных рядов и прогнозирование экономических показателей. – Харьков: Харьковский университет имени В.И.Каразина, 1999. – 229 с.
4. Клебанова Т.С., Иванов В.В., Дубровина Н.А. Методы прогнозирования: учебное пособие. – Х., 2002. – 372 с.
5. Комашко О.В. Практикум з прогнозування. – К.: 2000.
6. Лук'яненко І.Г., Городніченко Ю.О. Сучасні економетричні методи у фінансах. Навчальний посібник. – К.: Літера ЛТД, 2002. – 352 с.
7. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. – М.: «Дело», 2004.
8. Ставицький А.В., Руденський Р.А., Иванов В.В. „Часові ряди” у зб. Прикладная экономика (для магистров). – Т. 2. – Базовые модули. – Донецк, 2004. – С. 141–272.
9. Холден К., Піл Д., Томпсон Дж. Економічне прогнозування. Вступ. – К.:Інформтехніка, 1996. – 216 с.
10. Черняк О.І., Ставицький А.В. Динамічна економетрика. – К.: КВІЦ, 2000. – 120 с.
11. Черняк О.І., Ставицький А.В. Досвід розробки інформаційної системи прогнозування прибутків та збитків підприємств України // Проблемы развития внешнеэкономических связей и привлечения иностранных инвестиций: региональный аспект. – 2003. – Т. 3. – С. 704-706.
12. Черняк О.І., Ставицький А.В. Навчально-методичний комплекс з курсу „Часові ряди”. – Київ, 2004, 24 с.
13. Chatfield C. The Analysis of Time Series: An Introduction. – London, 1996. – 283 p.
14. Diggle P. J. Time Series: A Biostatistical Introduction. – Oxford, 1990. – 257 p.
15. Econometric Views 4.0 User's Guide. Quantitative Micro Software, LLC // [www.eviews.com](http://www.eviews.com)
16. Green, W.H.(2003), Econometric Analysis, 5th edition, Prentice Hall.
17. Mills, T.C. The Econometric Modelling of Financial Time Series. Cambridge University Press, 1999.
18. [www.andriystav.kiev.ua](http://www.andriystav.kiev.ua)

**Для нотаток:**