

Додаток А

Дослідження часового ряду даних на наявність сезонності

Учбовий приклад №1. Дослідження ВВП України.

Вхідні дані ВВП України в фактичних цінах, (млн. грн.) за період з 2002 року 1-го кварталу по 2007 рік 4-й квартал. Дані взяті з сайту державного комітету статистики <http://www.ukrstat.gov.ua/>.

Таблиця 1

Дані ВВП України за період з 2002 року 1-го кварталу по 2007 рік 4-й квартал.

t – номер виміру	Рік, квартал	y – ВВП (млн. грн.)
1	2002.1	44132
2	2002.2	50117
3	2002.3	65067
4	2002.4	66494
5	2003.1	52583
6	2003.2	60798
7	2003.3	75812
8	2003.4	78151
9	2004.1	66981
10	2004.2	78607
11	2004.3	99405
12	2004.4	100120
13	2005.1	88104
14	2005.2	101707
15	2005.3	122861
16	2005.4	128780
17	2006.1	105423
18	2006.2	124116
19	2006.3	150434
20	2006.4	157694
21	2007.1	133108
22	2007.2	161388
23	2007.3	195624
24	2007.4	222825

Крок-1. Створіть робочий файл, причому вкажіть квартальний (**Quarterly**), а діапазон значень від 2002:1 до 2007:4, рис. А.1.

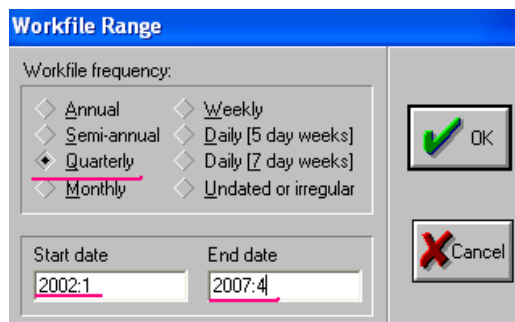


Рис. А.1

Крок-2. Завантажте учбовий файл **LR3_GDP-UA.txt** в систему Eviews, надайте часовому ряду ім'я “у”.

Крок-3. Побудуйте графік ряду у.

На рис. А.2 побудований графік для часового ряду ВВП. Візуально за графіком видно, що присутній тренд, який описується поліномом другого порядку.

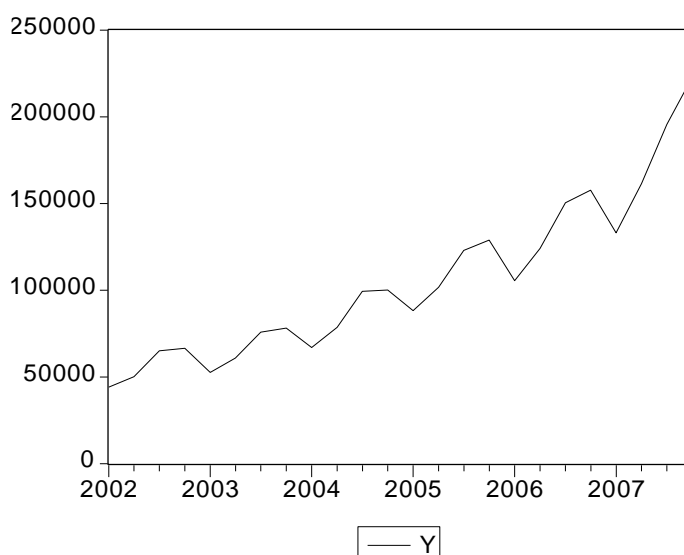


Рис. А.2

Визначення рівняння тренду

Крок-4. Завантажте файл LR3_time.txt в систему Eviews, надайте цьому ряду ім'я “t”. Цей ряд зберігає звичайні дискретні значення 1, 2, 3, 4, ..., 23, 24

Крок-5. Обчисліть коефіцієнти рівняння $y = a_0 + a_1 \cdot t + a_2 \cdot t^2$.

Для цього в командному режимі системи Eviews виконайте наступну команду

ls y=c(1)+c(2)*t+c(3)*t*t

Отримавши коефіцієнти моделі $a_0 = 54136,24$; $a_1 = 123,3399$; $a_2 = 243,6856$ можна описати тренд рівнянням $y(t) = 54136,24 + 123,3399 \cdot t + 243,6856 \cdot t^2$.

Крок-6. Створить новий часовий ряд з ім'ям “z” та запишіть в нього точки тренду. Для цього виконайте команди

```
series z  
z=c(1)+c(2)*t+c(3)*t*t  
plot y z
```

На рис. А.3 наведені графіки ряду “y” та його тренду “z”.

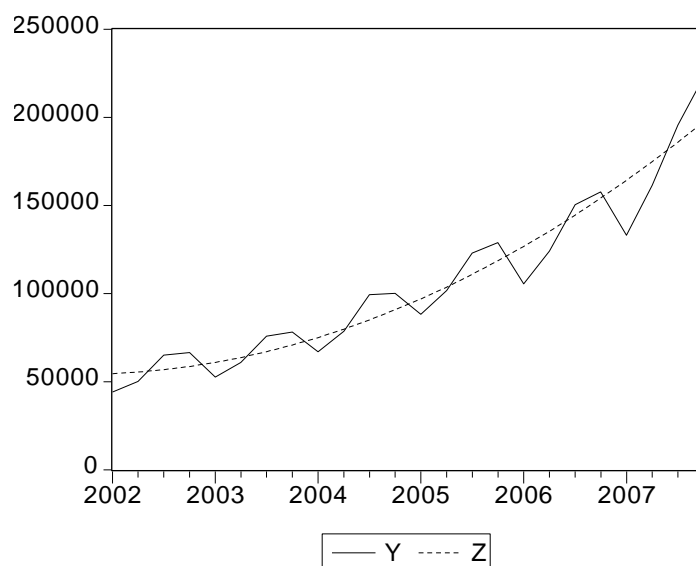


Рис. А.3

Крок-7. Видалення тренду “z” з часового ряду даних “y”.

Для цього створіть новий ряд “x”, та з ряду “y” видаліть ряд “z”. Для цього виконайте команди

```
series x  
x=y-z  
plot x
```

На рис. А.4 наведений графік для “x”.

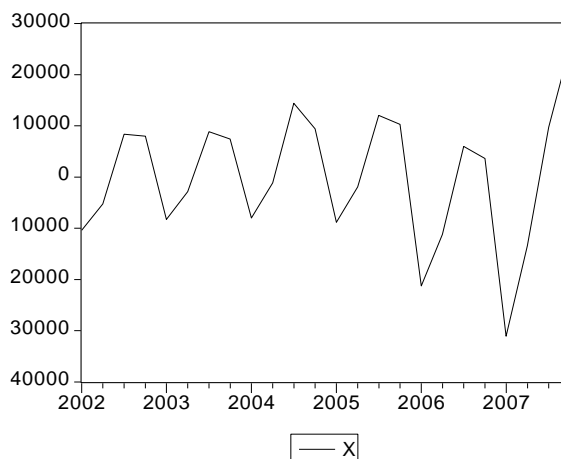


Рис. А.4 Часовий ряд даних ВВП з видаленим трендом (“ x ” = “ y ” – “ z ”)

Дослідження сезонності

Крок-8. Дослідження сезонного ефекту (сезонності) можна виконати двома способами візуально та по аналізу АКФ (ЧАКФ).

8.1. Візуально. Візуально по графіку на рис. А.4 видно, що функція має періодичний коливальний характер, причому період коливань дорівнює чотирьом.

8.2. Аналіз АКФ (ЧАКФ). Побудуйте для ряду “ x ” корелограму.

8.2.1. 1-й спосіб побудови корелограми. View→Correlogram.

В робочому файлі (**Workfile**) подвійним кліком лівої кнопки миші відкрийте вікно ряду “ x ”. У вікні, що з’явилося оберіть в меню закладки **View**, а в ній пункт **Correlogram**. Після цього з’явиться вікно, рис. А.6.

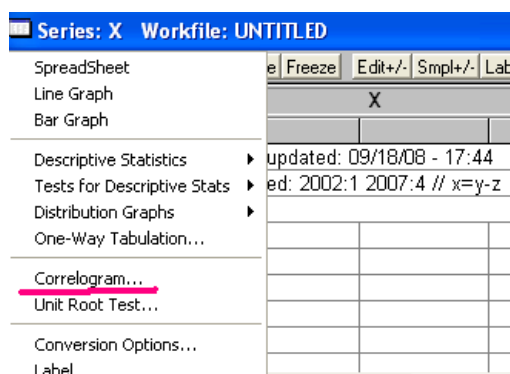


Рис. А.5

Опис опцій вікна **Correlogram Specification**, рис. А.6:

Level – побудова АКФ та ЧАКФ для ряду даних без змін.

1-st difference – побудова АКФ та ЧАКФ для ряду даних перших різниць. Від кожного k -го значення ряду буде відняте попереднє $k-1$ значення, тобто спочатку буде виконана операція $y(k)=y(k)-y(k-1)$.

2-nd difference – побудова АКФ та ЧАКФ для ряду даних перших різниць. Від кожного k -го значення ряду буде відняте попереднє $k-2$ значення, тобто спочатку буде виконана операція $y(k)=y(k)-y(k-2)$.

Lags to include – кількість значень ряду, для яких буде виведений результат.

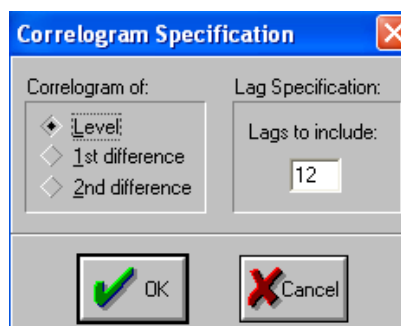


Рис. А.6

8.2.2. 2-й спосіб побудови корелограми. В командному режимі виконайте команду *x.correl*.

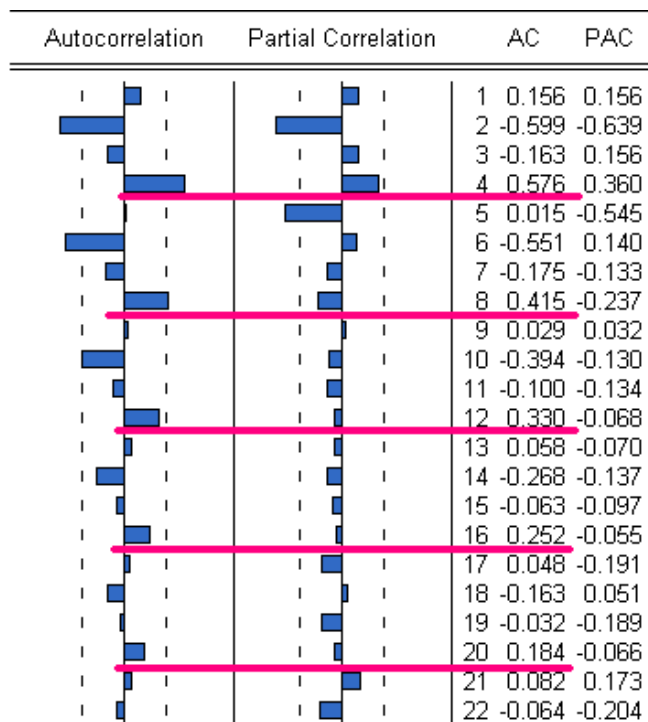


Рис. А.7

Як можна побачити з рис. А.7 дивлячись на значення АКФ (стовпчики **Autocorrelation** / **AC**) на 4, 8, 12, 16, 20 лагах, тобто з періодичністю чотири, проявляється сезонність. В ідеальному випадку дивлячись на значення ЧАКФ (стовпчики **Partial Correlation** / **PAC**) також можна побачити періодичність, але в даному випадку така ситуація не спостерігається.

Найбільш кмітливі студенти відмітять що ситуацію на рис. 8 можна трактувати також як прояв сезонного ефекту на 2, 6, 10, 14, 18, 22 лагах (рис. А.8). Але в цьому випадку періодичністю знову дорівнює чотирьом, тобто та ж сама ситуація але “в профіль”.

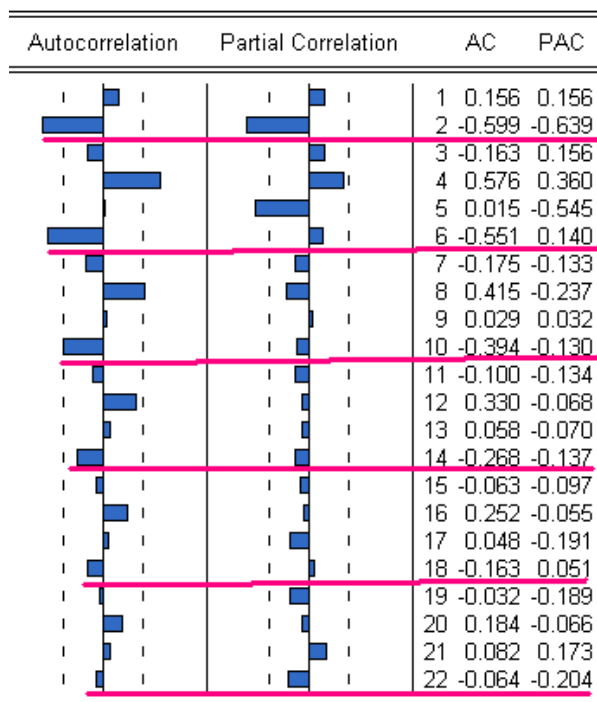


Рис. А.8

Крок-9. Зменшення впливу сезонного ефекту. На попередньому кроці-8 було визначено, що сезонність має періодичний характер, яка проявляється на кожному 4-му лагу. Щоб зменшити сезонний вплив, необхідно зменшити дисперсію процесу виконайте команди. Для цього створіть новий ряд з ім'ям “s” та виконайте команд

series s

s=x-x(-4)

Ще раз відмітимо, що ми не видаляємо сезонність, а прагнемо стабілізувати процес, привівши його до стаціонарного, шляхом зменшення дисперсії процесу.

На рис. А.9 наведений графік ряду “s”.

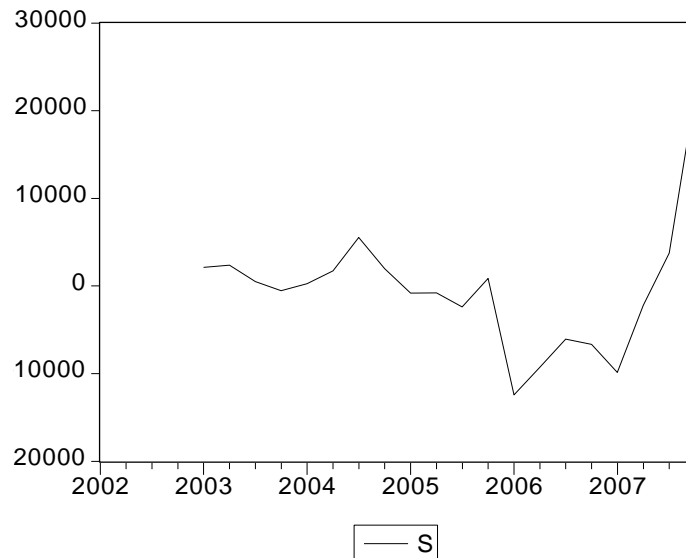


Рис. А.9

Крок-10. Обчисліть АКФ та ЧАКФ для ряду “s”, рис. А.10. Аналізуючи значення ЧАКФ видно що ряд “s” описується $AR(1)$. Це означає що для сезонний ефект можна описати рівнянням виду $x(k) = a_0 + a_1 \cdot x(k-1) + a_4 \cdot x(k-4)$, інформація про складову $a_0 + a_1 \cdot x(k-1)$ визначена тільки що на кроці-10, а про складову $a_4 \cdot x(k-4)$ на кроці-8.

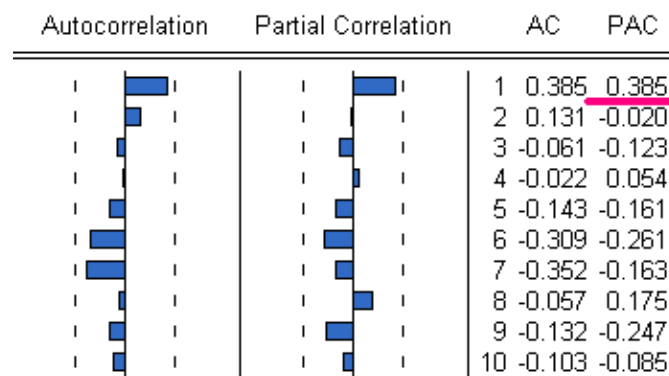


Рис. А.10

Виконавши команду

`ls x=c(1)+c(2)*x(-1)+c(3)*x(-4)`

знайдемо $a_0 = -501,4994$; $a_1 = 0,0937$; $a_4 = 1,1732$.

Висновок:

Отримавши на кроці-5 рівняння тренду, та на кроці-10 рівняння сезонної складової можемо записати рівняння моделі процесу наступним чином

$y(t) = z(t) + x(t),$	(A.1)
-----------------------	-------

де

$z(t) = 54136,24 + 123,3399 \cdot t + 243,6856 \cdot t^2,$	(A.2)
--	-------

$x(t) = -501,4994 + 0,0937 \cdot x(t-1) + 1,1732 \cdot x(t-4).$	(A.3)
---	-------

Підставивши значення в виведені рівняння А.1-А.3 можна побудувати прогноз на майбутній період часу:

$$z(t = 25) = 54136,24 + 123,3399 \cdot 25 + 243,6856 \cdot 25^2 \approx 209523,$$

$$x(t = 25) = -501,4994 + 0,0937 \cdot 25365,72 + 1,1732 \cdot (-31083,71) \approx -34592,$$

$$y(t = 25) = z(t = 25) + x(t = 25) \approx 209523 - 345912 = 174931.$$

Таблиця А.2

Дані для прогнозування ВВП

t	$y(t)$	$z(t)$	$x(t)$
1	44132	54503.27	-10371.27
2	50117	55357.66	-5240.665
3	65067	56699.43	8367.568
4	66494	58528.57	7965.429
5	52583	60845.08	-8262.081
6	60798	63648.96	-2850.962
7	75812	66940.21	8871.786
8	78151	70718.84	7432.163
9	66981	74984.83	-8003.832
10	78607	79738.20	-1131.197
11	99405	84978.93	14426.07
12	100120	90707.04	9412.959
13	88104	96922.52	-8818.520
14	101707	103625.4	-1918.370
15	122861	110815.6	12045.41
16	128780	118493.2	10286.82
17	105423	126658.1	-21235.15
18	124116	135310.5	-11194.48
19	150434	144450.2	5983.815
20	157694	154077.3	3616.739
21	133108	164191.7	-31083.71
22	161388	174793.5	-13405.53
23	195624	185882.7	9741.283
24	222825	197459.3	25365.72
25	174931	209523	-34592

За даними державного комітету статистики ВВП України в першому кварталі 2008 року склало 141084 млн. грн., а ми отримали прогнозне значення 174931. Такий великий розбіг між прогнозним та реальним значенням можна пояснити наступними факторами:

1. рівняння А.3 розходиться, тому що $a_1 + a_2 = 0.0937 + 1.1732 > 1$;

2. окрім дослідження самого ВВП необхідно включити в праву частину інші пояснюючі змінні - регресори. До цих регресорів можуть входити макроекономічні показники та індикатори: дані щодо зовнішньоекономічної діяльності, розмір податкових надходжень, інвестицій, інфляції та інші.