

НАДБУДОВА MS EXCEL ДЛЯ ДОСЛІДЖЕННЯ ЧАСОВИХ РЯДІВ

Анотація. Виконано розробку надбудови Excel (XLA), яка може бути використана для аналізу даних в електронних таблицях. Модулі надбудови надають можливість перегляду даних, експоненціального згладжування, нелінійного перетворення Бокса-Кокса та згладжування даних з використанням алгоритмів цифрової фільтрації. Алгоритми реалізовані в окремих вікнах, надбудова має зручний інтерфейс. Проведено перевірку роботи модулів.

Ключові слова: MS Excel, надбудова, часові ряди, експоненційне згладжування, перетворення Бокса-Кокса, КІХ-фільтри.

Постановка проблеми

Табличний процесор Microsoft Excel широко використовується для вирішення різноманітних задач, в тому числі і при дослідженні часових рядів та прогнозуванні. Функції, які вбудовані в MS Excel, можуть бути також поширені за допомогою модулів, які надаються в табличний процесор у вигляді нових надбудов [1]. Excel, як і інші програмні пакети, використовує роботу з надбудовами дуже широко, тому що надбудови поширюють можливості основних пакетів.

Актуальність розробки такого програмного забезпечення обумовлена, з одного боку, розповсюдженням MS Excel і зручністю його використання, а з другого - відсутністю низки вбудованих функцій, які необхідні для аналізу часових рядів.

Формулювання цілей статті

Метою даної роботи було поширення можливостей електронних таблиць MS Excel при вирішенні задач аналізу часових рядів, а саме, додавання модулів експоненційного згладжування, перетворення Бокса-Кокса та згладжування цифровим фільтром з кінцевою імпульсною характеристикою (КІХ) у вигляді надбудови Excel.

Основна частина

Надбудова MS Excel, яку розроблено, містить кілька модулів для аналізу часових рядів. На рисунку 1 представлено діаграму UML, на який показані алгоритми модулів.

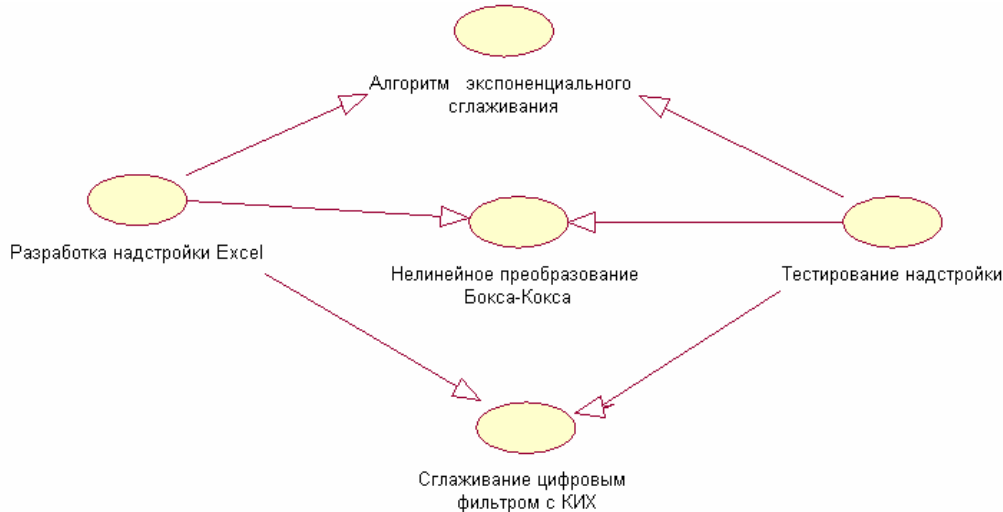


Рисунок 1 – Діаграма варіантів використання для надбудови Excel

Проведено розробку модулів експоненційного згладжування, нелінійного перетворення Бокса-Кокса та згладжування цифровим фільтром. При розробці надбудови використано C++ Builder, Visual Basic for Application (VBA), MS Excel.

У роботі [2] представлені попередні результати розробки надбудови MS Excel. Продемонстровано структуру надбудови як програмного засобу та зовнішній інтерфейс. Наведено приклади аналізу часових рядів методом експоненційного згладжування, та перетворенням Бокса-Кокса. В наступній роботі для дослідження часових рядів доданий модуль згладжування з використанням КІХ- фільтрів.

Розглянемо згладжування шляхом лінійної фільтрації даних цифровим фільтром з кінцевою імпульсною характеристикою. Лінійну фільтрацію доцільно проводити також і у частотній області [3,4], де можливо виявити періодичні коливання даних.

Нехай $X=\{x_1, x_2, x_3 \dots x_N\}$ – часовий ряд (сигнал), який надано, $Y=\{y_1, y_2, y_3 \dots y_N\}$ – часовий ряд (сигнал) після згладжування.

Згортка сигналу має вигляд

$$y(n) = \sum_{m=0}^{\infty} h(n-m)x(m) = \sum_{m=0}^{\infty} h(m)x(n-m),$$

де m - затримка послідовності, $h(n)$ – імпульсна характеристика (ІХ).

$H(z)$ — z -зображення ІХ, $h(n)$ — обчислюється за формулою

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n}.$$

Передавальна функція $H(z)$ дорівнює

$$H(z) = Y(z)/X(z),$$

де $X(z)$, $Y(z)$, — z - зображення сигналу до- і після обробки.

Нехай $H(e^{j\hat{\omega}})$ - Фур'є - зображення ІХ $h(n)$, де $\hat{\omega}$ — нормована частота.

Для частотної характеристики $H(e^{j\hat{\omega}})$ і передавальної функції $H(z)$ справедливо співвідношення

$$H(e^{j\hat{\omega}}) = H(z) \Big|_{z=e^{j\hat{\omega}}}$$

Таким чином, маємо змогу проводити аналіз даних у частотній області.

У модулі надбудови реалізовано кілька варіантів згладжування сигналу. До простих алгоритмів згладжування належать 15 і 21 точкова формула згладжування по Спенсеру. Використані формули згладжування за методом найменших квадратів, які отримані при проведенні параболи через кілька рівновіддалених значень $X(k)$, з подальшим використанням в якості згладженої величини значення параболи в середній точці (фільтри Савицького-Голея).



Рисунок 2 – Результати цифрової фільтрації

На рисунку 2 показані результати згладжування та згортки з кінцевою імпульсною характеристикою цифрового фільтра усереднення по 12 точкам відповідно формулі

$$Y(n) = (X(n-6) + X(n-5) + X(n-4) + X(n-3) + X(n-2) + X(n-1) + X(n) + X(n+1) + X(n+2) + X(n+3) + X(n+4) + X(n+5)) / 12.$$

Результатом лінійної фільтрації (згортки), є послідовність чисел, число яких дорівнює

$$N_CONVOL = N_DATA + N_FIR - 1.$$

де N_DATA - число даних часового ряду, N_FIR - довжина кінцевої імпульсної характеристики цифрового фільтра.

Перевірку алгоритмів, які запропоновано, проведено з використанням часового ряду, що характеризує чисельність пасажирів, які використовували повітряний транспорт із січня 1949 року по грудень 1961 року (за 12 років щомісячно) у США. Дані представлені у вигляді стовпця чисел і показують кількість тисяч пасажирів на місяць. Усього 144 точки. Дані взяті з файлу `airpass.dat` [5] і є відомим тестовим сигналом.

Алгоритми експоненційного згладжування та нелінійного перетворення Бокса-Кокса протестовані з використанням широко відомого пакету ITSM 2000 [5, 6], згладжування цифровим фільтром з кінцевою імпульсною характеристикою – з використанням MatLab [7].

На рисунках 3 і 4 показані види екранів додатку ITSM2000 [9] та надбудови Excel для алгоритму експоненційного згладжування. Графіки на рисунках – ідентичні.

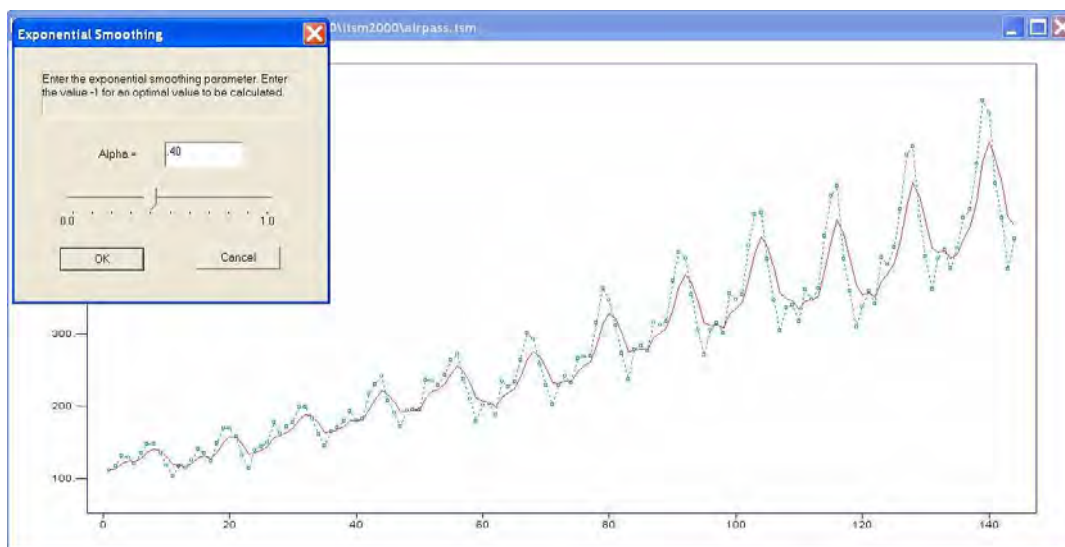


Рисунок 3 - Експоненціальне згладжування в ITSM 2000

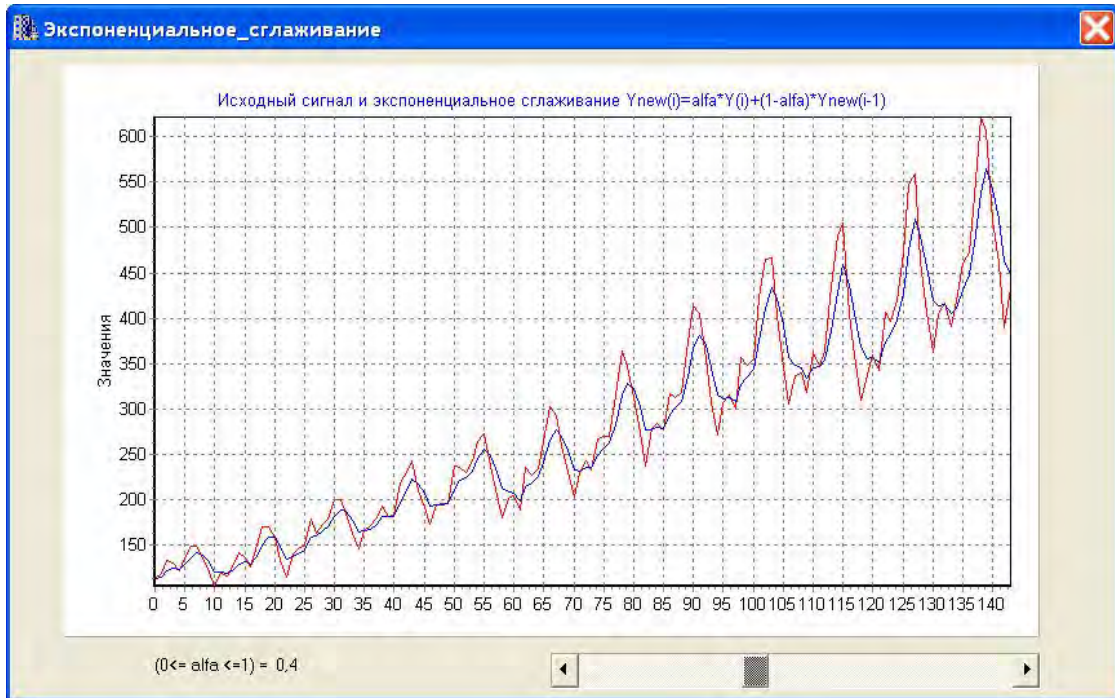


Рисунок 4 - Экспоненціальне згладжування у «Надбудові»

На рисунках 5 і 6 показані види екранів додатку ITSM2000 [9] та надбудови Excel для алгоритму нелінійного перетворення Бокса-Кокса. Результати у графічному вигляді схожі.

Для тестування алгоритму згладжування цифровим фільтром використано MatLab. Використані функції БПФ (Швидкого перетворення Фур'є – fft) та згортки двох послідовностей ($conv$). Отримана якісна збіжність результатів (Таблиця 1).

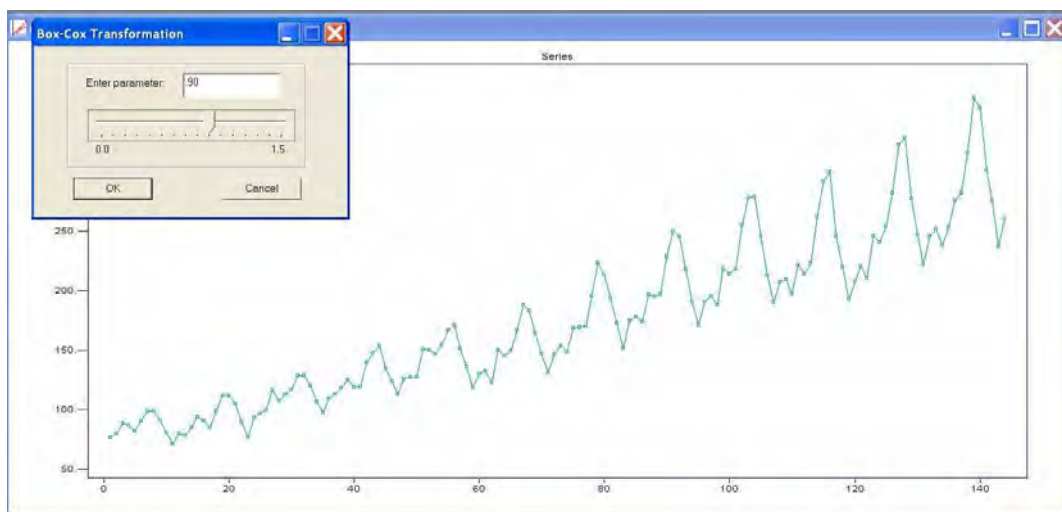


Рисунок 5 - Нелінійне перетворення Бокса-Кокса в ITSM 2000

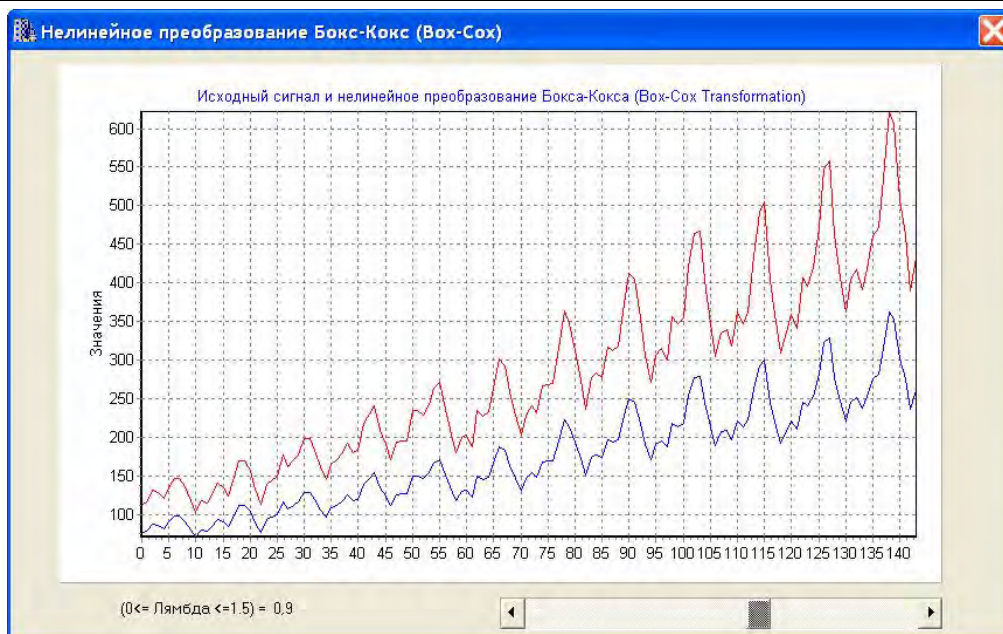


Рисунок 6 - Нелінійне перетворення Бокса-Кокса в «Надбудові»

Порівняльний аналіз тестування алгоритмів експоненціального згладжування, нелінійного перетворення Бокса-Кокса та алгоритму згладжування цифровим фільтром представлений у таблиці 1. Якісну оцінку результатів показує значення Δ - абсолютна похибка.

Висновки

Розроблена надбудова Excel (XLA), для обробки часових рядів в електронних таблицях MS Excel.

Модулі надбудови надають можливість перегляду даних та їх аналізу. Алгоритми реалізовані в окремих вікнах і мають зручний інтерфейс. За допомогою функцій надбудови перетворені (згладжені) дані часового ряду експортуються до комірок Excel.

Реалізовані функції згладжування даних методами експоненційного згладжування, нелінійного перетворення Бокса-Кокса та цифрової фільтрації.

При цифровій фільтрації передбачені декілька методів усереднення, в тому числі фільтрами Савицького-Галея і фільтрами Спенсера (15 і 21 точкове усереднення). Показано, що при лінійній фільтрації зручно використовувати частотну область, в якій можливо виявити періодичні коливання даних.

Проведено перевірку модулів надбудови Excel. Доведено, що модулі роботають якісно.

Результаты порівняльного аналізу функцій надбудови
та відомих пакетів

Сигнал	Надбу- дова	ITSM 2000	Δ	Надбу- дова	ITSM 2000		Надбудова	Mat Lab	
air pass	Expsmooth $\alpha=0,4$		Δ exp sm	BoxCox $\lambda=0,9$		Δ Box Cox	KIX фільтр (центральна частина)		Δ KIX
112	112,00	112,00	0,00	76,52	76,52	0,00E+00	74,58	74,58	5,97E-13
118	114,40	114,40	0,00	80,26	80,26	0,00E+00	86,92	86,92	1,56E-12
132	121,44	121,44	0,00	88,90	88,90	1,56E-13	98,25	98,25	-4,41E-13
129	124,46	124,46	0,00	87,05	87,05	0,00E+00	108,17	108,17	2,42E-13
121	123,08	123,08	0,00	82,12	82,12	0,00E+00	116,83	116,83	6,68E-13
.....
472	447,33	447,33	0,00	282,23	282,23	6,25E-13	473,92	473,92	-1,31E-12
535	482,40	482,40	0,00	316,05	316,05	5,12E-13	476,17	476,17	-3,01E-12
622	538,24	538,24	0,00	362,11	362,11	0,00E+00	441,42	441,42	0,00E+00
606	565,34	565,34	0,00	353,69	353,69	0,00E+00	408,83	408,83	-1,42E-12
508	542,41	542,41	0,00	301,60	301,60	7,96E-13	373,92	373,92	-1,65E-12
461	509,84	509,84	0,00	276,28	276,28	5,68E-13	335,50	335,50	-3,41E-12
390	461,91	461,91	0,00	237,51	237,51	6,25E-13	296,17	296,17	-6,25E-13
432	449,94	449,94	0,00	260,52	260,52	0,00E+00	251,58	251,58	-2,10E-12

ЛІТЕРАТУРА

1. Гайдышев И.П. Решение научных и инженерных задач средствами Excel, VBA и C/C++. –СПб.: БХВ – Петербург, 2004. – 512с.
2. Дорош Н.Л., Храпач Ю.О. Результаты розробки надбудови MS Excel для обробки часових рядів // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 2 (97). – Дніпропетровськ, 2015. – С. 8-12.
3. Хемминг Р.В. Цифровые фильтры. Пер. с англ. /Под ред. А. М. Трахтмана. – М. – Сов. радио, 1980.- 224с.
4. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. Основные методы. Пер. с англ. /Под ред. В. И. Хохлова. – М. – Мир, 1982. – 428с.
5. Peter J. Brockwell, Richard A. Davis. ITSM for Windows. A User's Guide to Time Series Modelling and Forecasting. – Springer-Verlag New York, Ins., 1994.
6. ITSM 2000 Professional Version 6.0, developed by Peter J. Brockwell and Richard A. Davis, B&D Enterprises, Inc., Copyright 1999.
7. Цифровая обработка сигналов и MATLAB: учеб. пособие / А. И. Солонина, Д.М. Клионский, Т.В. Меркучева, С.Н. Перов. — СПб.: БХВ-Петербург, 2013. — 512 с.