

УДК 519.6

АПРОКСИМАЦІЯ ТА ПРОГНОЗ МАЙЖЕ ПЕРІОДИЧНИХ ПРОЦЕСІВ

О.М. Дрєєв , З.Ю. Філер

Рассмотрена последовательная аппроксимация методом наименьших квадратов тригонометрическим полиномом сигналов, которые содержат некрратные гармоники, с целью экстраполяции. Такие сигналы не являются периодическими и для прогнозирования их поведения стандартные Фурье и вейвлет-преобразования непригодны. Приведено описание алгоритма аппроксимации, выведена вероятностная оценка ошибки экстраполяции.

In the article successive approximation by a least-squares method is considered by the trigonometric polynomial of signals which contain aliquant accordions, with the purpose of extrapolation. Such signals are not periodic and for prognostication of their conduct standard Fourier and wavlet transformations are useless. Description of algorithm of approximation is resulted, estimation of error of extrapolation is shown out.

При традиційному проведенні гармонічного аналізу функція, що досліджується, вважається періодичною і розглядаються експериментальні дані на проміжку одного періоду [2, 26 с.]. Завдяки таким припущенням виконання екстраполяції для прогнозу поведінки систем стає неможливим. Для практичного використання в прогнозуванні плин природного процесу й призначена регресія (розклад) функції, що досліджується в тригонометричний ряд з некрратними гармоніками.

Задача полягає в апроксимації сигналу, заданого значеннями на (можливо нерівномірній) сітці часу неперіодичних коливних процесів, що містять в собі некрратні гармоніки. Розглядаємо наближення експериментальних даних функцією $g(t) = a_0 + a_1 \sin(\omega t) + a_2 \cos(\omega t)$, де a_0 , a_1 , a_2 , ω – невідомі коефіцієнти.

Дано: $f(t_i)$ – значення функції що досліджується в моменти часу t_i , де $i=1..N$, N – натуральне число; потрібно визначити коефіцієнти a_0 , a_1 , a_2 , ω так щоб

$$S(\omega) = \frac{1}{N} \sum_{i=1}^N (a_0 + a_1 \sin(\omega t_i) + a_2 \cos(\omega t_i) - f(t_i))^2 = \min . \quad (1)$$

Для знаходження a_0 , a_1 , a_2 використовується метод найменших квадратів.

Знайдемо частинні похідні функції $S(\omega)$ та прирівняємо їх до нуля, утворюючи систему рівнянь, розв'язок якої дає мінімум середньоквадратичного відхилення [3]:

$$\begin{aligned} a_0 + a_1 \overline{\sin(\omega t)} + a_2 \overline{\cos(\omega t)} - 2 \overline{f(t)} &= 0, \\ a_0 \overline{\sin(\omega t)} + a_1 \overline{\sin^2(\omega t)} + a_2 \overline{\sin(\omega t) \cos(\omega t)} - \overline{f(t) \sin(\omega t)} &= 0, \\ a_0 \overline{\cos(\omega t)} + a_1 \overline{\sin(\omega t) \cos(\omega t)} + a_2 \overline{\cos^2(\omega t)} - \overline{f(t) \cos(\omega t)} &= 0. \end{aligned} \quad (2)$$

Тут $\frac{1}{N} \sum_i v(t_i) = \overline{v(t)}$, де N – кількість доданків, i – індекс сумування.

Розв'язком системи рівнянь (2) є значення коефіцієнтів a_0, a_1, a_2 . Це дає можливість визначити функцію, яка має мінімальне середнє квадратичне відхилення при заданій частоті ω . За допомогою чисельного методу шукається мінімум функції середнього квадратичного відхилення від частоти $S(\omega)$. Цей мінімум буде найкращим середньоквадратичним наближенням коливною функцією на проміжку дослідження експериментальних даних [1, 64с.].

Пошук величини ω обмежується частотою Котельникова-Найквіста: при дискретизації сигналу з кроком в часі h , сигнал можна відновити однозначно, якщо він не містить складові частоти, вищої за $\omega = \pi/(2h)$. Теоретично нижня границя частот обмежена нулем, але для практичних розрахунків нижня межа частот обмежується трьома - десятьма проміжками дослідження сигналу.

Залежність локальних мінімумів од довжини проміжку аналізу сигналу для $\sin(t)$

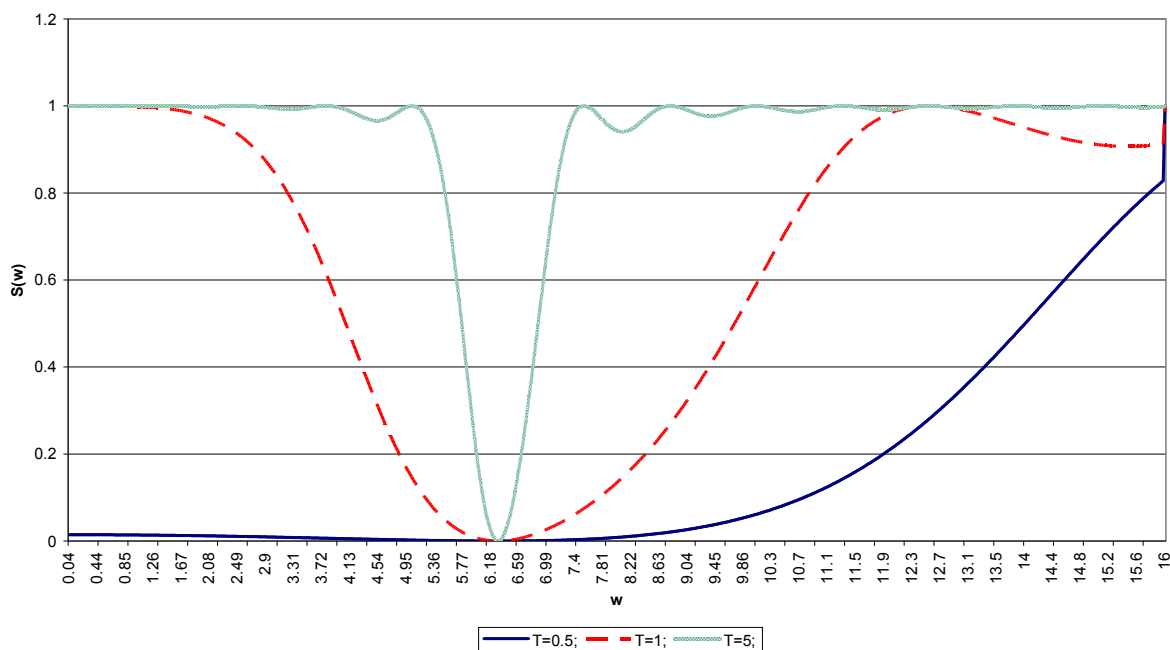


Рис. 1

Функція $S(\omega)$ має коливний характер (рис. 1), з локальними мінімумами в частотах базису тригонометричних функцій Фур'є (приклад побудовано для аналізу функції $y(t)=\sin(t)$). Визначити оціночні значення локальних мінімумів ω_k можна з часу спостереження сигналу $T=t_n-t_0$. У базис Фур'є входять кратні частоти до періоду спостереження, тому $\omega_k = 2\pi k/T$, де k – натуральне число. З рисунку 2 робиться висновок, що кількість точок

розбиття (дискретизація сигналу) мало впливає на положення локальних мінімумів, але вносить спотворення зміщення частоти Найквіста ω_H .

Знаходиться мінімальне значення таблично отриманої $S(\omega)$, розбивши проміжок шуканих $[\omega_0; \omega_H]$ частот на рівномірні проміжки кількістю, більше ніж

$$N_\omega = 3 \frac{\omega_H - \omega_0}{2\pi} (t_n - t_0), \quad (3)$$

Подальше уточнення частоти проводиться довільним чисельним методом, наприклад дихотомією.

Залежність локальних мінімумів до кількості точок розбиття для $\sin(t)$

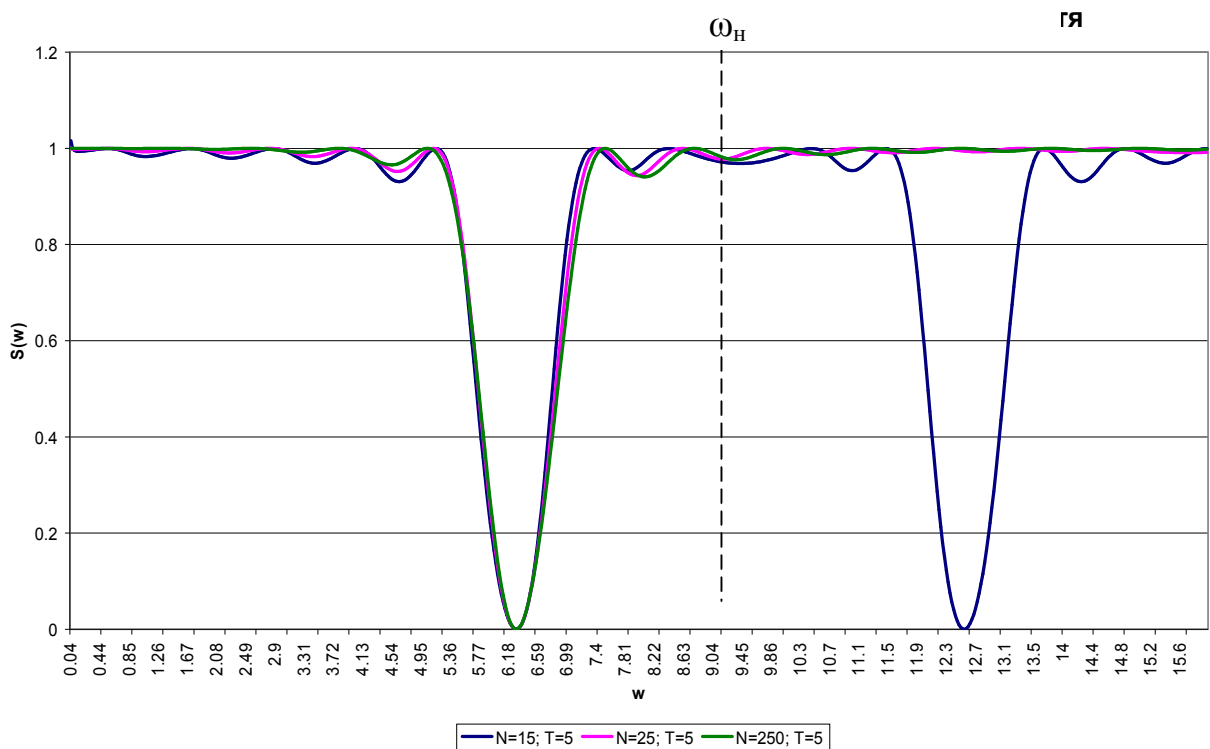


Рис. 2

Послідовне наближення функції тригонометричним поліномом дає апроксимацію:

$$f(t) \approx \sum_{i=1}^K g_i(t), \quad g_i(t) = a_{0,i} + a_{1,i} \sin(\omega_i t) + a_{2,i} \cos(\omega_i t). \quad (4)$$

Процедура описується алгоритмом:

1. Для функції $f(t)$ шукаємо наближення $g_1(t)$ знаходимо функцію відхилень $f_1(t) = f(t) - g_1(t)$.
2. Повторюємо п. 1 для наближення $f_1(t)$ і отримуємо $f_2(t)$, $g_2(t)$.
3. Дія повторюються до досягнення бажаної точності наближення (на кожному кроці маємо середнє квадратичне відхилення та набір відхилень наближення від реальних даних).

Виконання алгоритму дає гарантоване середньоквадратичне наближення (4) не гірше за ряд Фур'є, бо використовується ширший набір базисних функцій.

Головною метою розробки методу наближення тригонометричним многочленом є проведення прогнозування майже періодичних процесів, але для побудови прогнозів потрібно мати оцінку їх надійності.

Класично аналіз надійності наближення полягає в оцінці точності знаходження періоду та фази складових сигналу. При дослідженні ж майже періодичного сигналу, де період не стабільний, або сигнал складається з суми двох некратних частот, точне визначення періоду неможливе, бо він не існує. Автори дають ймовірність надійного прогнозу на заданому інтервалі.

Будемо вважати досліджуваний сигнал таким, що він містить у собі випадковий набір гармонічних коливань. Такий процес є ергодичним і для нього справедливе співвідношення:

$$\tau_p = \frac{1}{R_y(0)} \int_0^{\infty} |R_y(s)| ds. \quad (5)$$

Тут τ_p – порядок часу, на який має сенс прогнозування процесу, а границя

$$R_y(s) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(x)y(x-s)dx - \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(x)dx \right)^2$$

є кореляційною функцією досліджуваного сигналу.

Вважаємо відхилення $f_{k+1}(x_i)$ випадковою величиною, підпорядкованою нормальному закону розподілу з нульовим математичним сподіванням та середнім квадратичним відхиленням $I_k(\omega_k)$. Допустима норма b відхилення реального сигналу від прогнозу; ймовірність того, що ця норма не буде перевищена є

$$P_H \leq \left(\frac{1}{\sqrt{2\pi} \cdot I_k(\omega_k)} \int_{-b}^b \exp(-l^2 / (2I_k(\omega_k))) dl \right)^n,$$

де n – кількість кроків прогнозу; l – змінна інтегрування. В цьому випадку час прогнозування $t_H = hn$, з ймовірністю P_H є надійним для $[-b; b]$. В реальних розрахунках прогнозування має значно більшу точність, бо при оцінці ймовірності перевищення похибки прогнозування враховувався найгірший варіант розвитку подій.

Короткочасний прогноз чисел Вольфа (базові дані за 180 діб, прогноз на 60 діб)

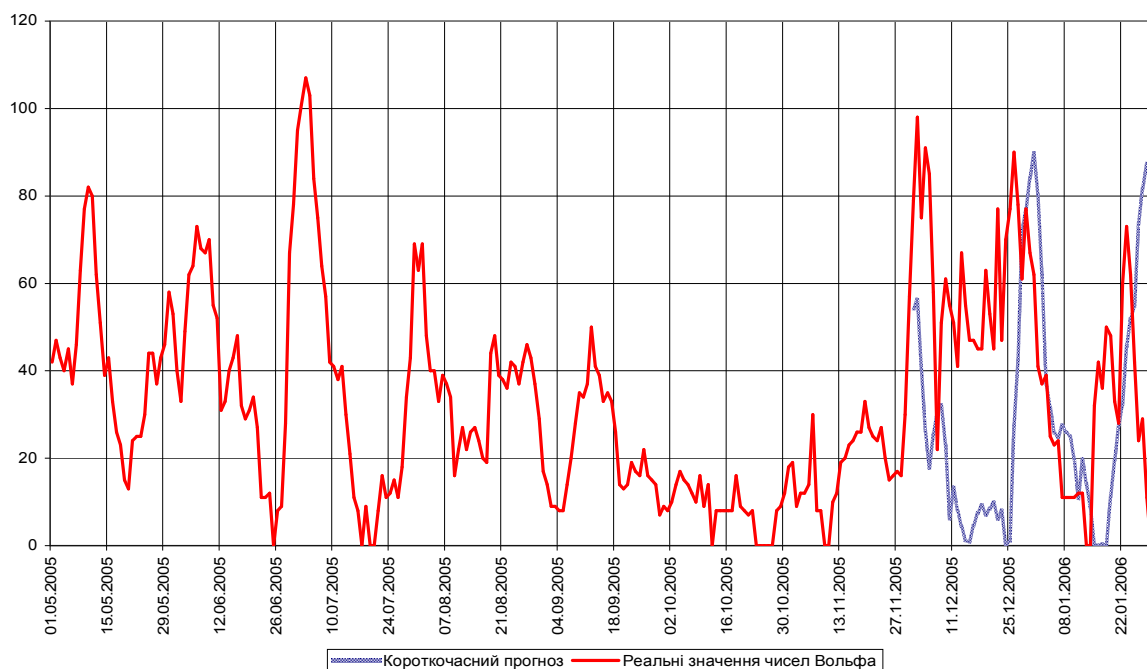


Рис. 3

Для перевірки можливостей методики прогнозування поведінки числових рядів, було проведено амплітудно-частотний аналіз середньорічних даних сонячної активності за 1749-2000 роки [4]. Частотний аналіз процесу зміни сонячної активності показує наявність на фоні майже періодичного основного сигналу стохастичної складової. Вона має певні частоти, але фаза коливань стохастичної складової та її амплітуда є випадковою величиною. За припущенням авторів збудниками стохастичних коливань є турбулентні збурення та вибухи на поверхні Сонця, а синхронізуючим фактором основних коливань сонячної активності є блукання центру Сонця навколо центру мас Сонячної системи під дією гравітації планет-гігантів. Ці чинники впливають на принципи побудови прогнозів зміни сонячної активності. Для довгострокових прогнозів врахування стохастичної складової є недоречним з огляду на нестабільність та малий, в середньому, вплив на довготривалі процеси. Тому для довгострокового прогнозування змін сонячної активності (СА) було взято середньорічні значення СА. Практична перевірка надійності прогнозування на довгостроковий період проводилася при побудові «прогнозу» в минуле до 1650 року включно. Отриманий прогноз досить точно відобразив дійсну картину змін сонячної активності – було чисельно отримано аномальне падіння сонячної активності, що відбувалося реально і відоме як «мінімум Маундера».

Короткочасні прогнози щоденних значень чисел Вольфа проводилися на строк від одного до трьох місяців, бо випадковість виникнення малопотужних коливань не прогнозується, а екстраполюється вже наявні збурення та частоти. Тому базою для проведення екстраполяції є період

спостережень близько шести місяців. В цих випадках прогнозування дає найкращі результати на проміжках стабільної зміни сонячної активності, при тенденціях зростання та спадання в середньому сонячної активності, прогнозування якісно відображає лише фазову характеристику, амплітудна ж характеристика в частині випадків не відповідала дійсності. Графіки змін сонячної активності та їх прогнози показані на рис.3 для довгострокового прогнозування та на рис.4 для короткострокового прогнозування.

“Передбачення” аномального зниження середньомісячних чисел Вольфа екстраполяцією в минуле

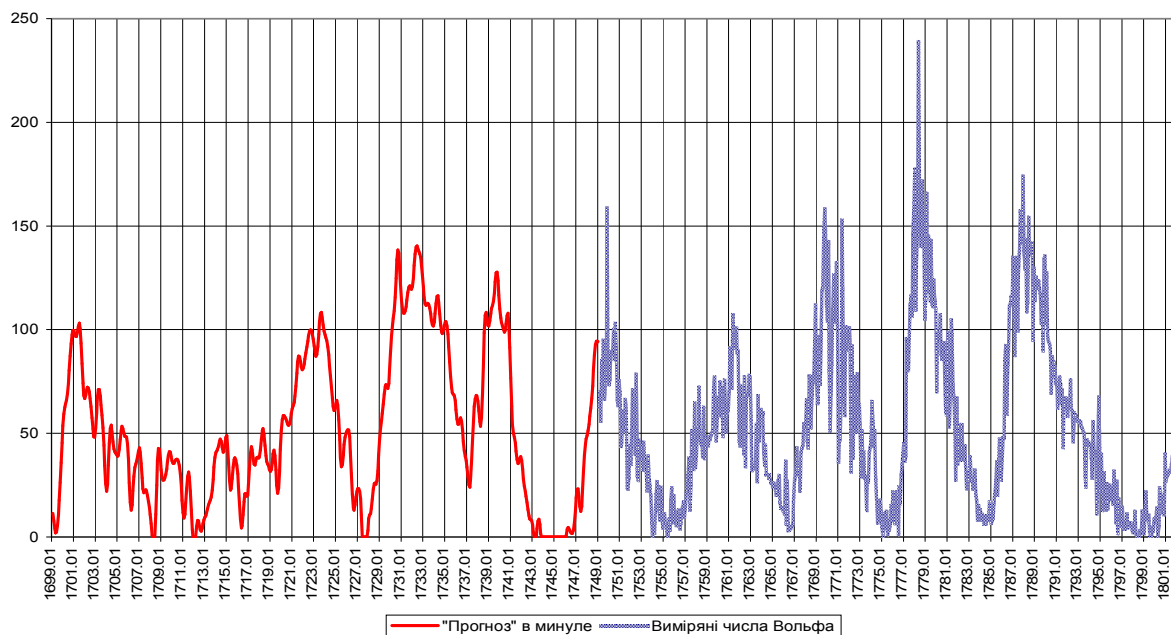


Рис. 4

Для підвищення точності прогнозів поведінки числових рядів автори планують визначити оптимальний проміжок інтерполяції сигналу та на його основі – оптимальний проміжок екстраполяції. Оцінка ймовірності можливих відхилень прогнозування є досить грубою, тому вона вимагає значного уточнення. Планується значно підвищити точність апроксимації та екстраполяції сигналів, врахувавши поточний коефіцієнт “затухання” коливного сигналу, що додає ще одну умову на дослідження сигналу – виміри сигналу, що досліджується, повинні бути рівномірними в часі.

Розроблені алгоритми застосовуються для прогнозування [5] стану здоров'я та психіки людей внаслідок щоденних змін СА, на 7-30 діб, який розміщується в ЗМІ, зокрема, в газеті “Моя Батьківщина/Моя Родина”, яка виходить у Харкові, та прогнозів метеоданих та врожаїв найближчих років.

Висновки

1. Розроблено чисельний алгоритм апроксимації сигналу не ортогональним базисом коливних функцій.
2. Побудована ймовірнісна оцінка похибки екстраполяції, що дозволяє прогнозувати числові послідовності. На практиці при дослідженні

змін сонячної активності було «прогнозовано» в минуле аномальний мінімум Маундера сонячної активності в 1650-1750 роках.

3. Визначено оптимальне початкове розбиття для уточнення основної частоти апроксимуючої функції.
4. Розроблено програму для проведення апроксимації та отримання прогнозів поведінки числових рядів.
5. Розроблена програма застосована до аналізу та прогнозуванню щоденних, середньомісячних та середньорічних значень чисел Вольфа – характеристики СА для розробки її прогнозів на різні терміни.
6. Досліджена кореляція середньорічних значень СА з кліматичними змінами, та середньомісячних – з погодними змінами.

БІБЛІОГРАФІЯ

1. Дреєв О.М., Філер З.Ю. Спектральний аналіз майже періодичних сигналів// Теорія та методика навчання математики, фізики, інформатики. Збірник наукових праць. Випуск 4. Том 3. - Кривий Ріг. Видавничий відділ НМетАУ, 2004. – С.64-68.
2. Блатер К. Вейвлет - анализ. Основы теории. – М.: Техносфера, 2004. - 274 с.
3. Марчук Г.И. Методы вычислительной математики. – М.: Наука, 1989, 608 с.
4. <http://www.spaceweather.com/>
5. Філер З.Ю. Яка погода буде наступного 2004 року в Космосі, на Землі та в Соціумі? – Кіровоград:Центр.-Укр. видав., 2004. – 24с.

*Кіровоградський державний педагогічний
університет ім. В.Винниченка*

Надійшло 18 березня 2006 р.