УДК 519.65
О.Б.Новікова
Національний авіаційний університет

ОПТИМІЗАЦІЯ ПРОПОРЦІЙ ФРАКТАЛЬНИХ СПЛАЙНІВ

У статті представлено автоматичний алгоритм пошуку оптимальних положень вузлів фрактального сплайна. Оцінка регресії здійснювалась за допомогою методу найменших квадратів. Розглянуто вплив близькості вузлів на якість апроксимації. Роботу алгоритму проілюстровано прикладом.
Ключові слова: *фрактальний сплайн, оптимізація, метод найменших квадратів, апроксимація.*

Актуальність

Важливість розміщення вузлів для звичайного сплайна, побудованого за МНК, давно відома. Доведено, що результати сплайн-апроксимації значно залежать від правильного розміщення вузлів сплайна на сітці. Розроблено багато методів детерміністичного та стохастичного пошуку (метод Гауса-Ньютона, випадковий пошук, генетичні алгоритми, метод відсікання кутів, DC-програмування, оптимізація Ліпшица тощо) [1,2,3].

Однак поза увагою залишились деякі інші типи сплайнів, зокрема фрактальні. Сьогодні ми можемо говорити, що велика кількість сигналів, з якими має справу наука, мають фрактальну природу. Класичні приклади: довжина узбережжя, сейсмічні коливання, людська мова, курси валют та ін. Фрактальні сплайни створені для роботи з такими сигналами, їх аналізу та прогнозування. Наступний приклад яскраво демонструє важливість оптимального розміщення вузлів для фрактальних сплайнів.

Рис.1 показує набір даних, згенерованих алгоритмом фрактального броунівського руху. Для апроксимації використано кубічний сплайн з 5 рівновіддаленими вузлами та фрактальний сплайн масштабу 2, що використовує цей кубічний сплайн як базовий. Середньоквадратична помилка для кубічного сплайну складає 2,0238, а для фрактального сплайну – 2,1857.

З цього можна зробити помилковий висновок, що звичайна сплайн-регресія краща за фрактальну у будь-якому випадку. Але той факт, що звичайний сплайн не враховує самоподібність, робить практично неможливим їх адаптацію до локальних особливостей фрактальних функцій. Однак фрактальний сплайн відрізняється від звичайного лише глибиною вкладених рівнів, а значить помилка, допущена при розміщенні вузлів на першому рівні, з кожним наступним масштабом буде збільшуватися. Тому оптимальне розміщення вузлів у фрактальному сплайні є необхідною умовою для отримання всіх переваг самоподібності.

Задачі дослідження:

- аналіз залежності між розміщенням вузлів фрактального сплайна та якістю апроксимації,
- пошук шляхів оптимізації розміщення вузлів для підвищення якості апроксимації за вибраним критерієм,
- модифікація методів оптимізації розміщення вузлів, розроблених для кубічних сплайнів, до застосування з фрактальними сплайнами,
- проведення комп'ютерного експерименту з апроксимації фрактальних даних та порівняння результатів до та після застосування алгоритму оптимізації вузлів.

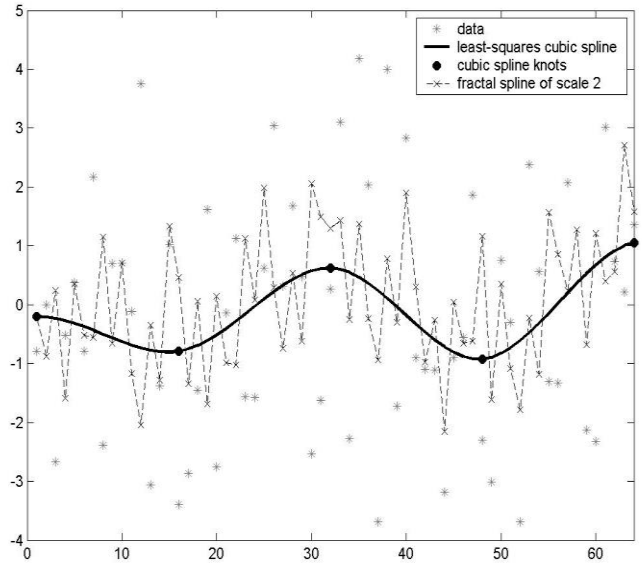

Рис. 1. Порівняння апроксимації кубічним сплайном та фрактальним кубічним сплайном масштабу 2.

Оптимізація розміщення вузлів фрактального сплайна

Якщо вхідний інтервал апроксимації розбити на однакові елементарні відрізки $\xi_n - \xi_{n-1} \equiv H$, то інформація про функцію, що апроксимується, буде рівномірно розподілена по всьому інтервалу. Але інтенсивність функції на інтервалі не обов’язково буде однаковою. Таким чином, оптимальний метод розбиття інтервалу апроксимації на відрізки має враховувати:

а) кількість вузлів повинна бути малою порівняно з кількістю точок,

б) чим вище щільність початкових точок, тим більше вузлів повинно бути розміщено в даній області,

в) вузли слід розміщувати на границях змін функції регресії.

Нехай маємо деякий фрактальний сигнал, який виміряно на інтервалі $a \le \tau \le b$ з кількістю точок n. Нехай ξ - набір вузлів, кожна точка якого $\xi_i, i = 1:r$, при чому $\xi_1 = a, \xi_{i+1} > \xi_i$ і $\xi_r = b$. ***Кубічним сплайном*** з вузлами ξ будемо називати функцію на інтервалі $[a,b]$, яка є поліномом третьої степені на кожному підінтервалі $[\xi_i, \xi_{i+1}]$ і має дві неперервні похідні на інтервалі $[a,b]$. Позначимо множину кубічних сплайнів, які можна побудувати на ξ, як S_ξ. За умови неперервності другої похідної кількість таких сплайнів дорівнює $(r+2)$.

Фрактальний сплайн – це функція, що складається з сплайн-функцій різного масштабу, що зберігають самоподібність [4, с. 50]. Під масштабом розуміють кількість вкладених рівнів самоподібних сплайнів. Для того, аби сплайн став фракталом, необхідно, щоб кожен із $(r-1)$ фрагментів також був сплайном, подібним до оригінального. Тоді сплайн на k-му масштабі буде складатися з $(r-1)^k$ фрагментів. Поділ кожного фрагменту сплайна зберігає пропорцію нульового масштабу. Кількість фрактальних сплайнів масштабу k, які можна побудувати на сплайнових базисах S_ξ, дорівнює $S_\xi^k = (r+2)^k$.

Задача полягає у відшуканні такого фрактального сплайну ${}^kFS$ масштабу k, який апроксимує дані з найменшою похибкою. Традиційно ця задача розв’язується за методом найменших квадратів (МНК). Але як ***критерій оптимізації*** також можуть використовуватись функції:

1. *Фрактальної розмірності*. Знайти криву, яка апроксимує (або інтерполює) набір вхідних точок і має найменшу фрактальну розмірність. Така необхідність може виникнути при аналізі деяких природних фрактальних об’єктів (морського узбережжя, зображення людських тканин і органів), адже поняття довжини, площі, кривизни тощо є некоректними для фракталів, їх вимірюють опосередковано через фрактальну розмірність.

2. *Інформації*. Завдяки самоподібності фрактали широко застосовуються в алгоритмах стиснення зображень та передачі інформації. Тому може постати задача знайти такий сигнал, який має задані характеристики та передає найбільше інформації.

У даній статті ми розглянемо оптимізацію вузлів фрактального сплайна за критерієм найменшої похибки.

Нехай $\hat{s}_f$ - шуканий фрактальний сплайн, який найкраще апроксимує вхідний сигнал y_j :

$$\hat{s}_f \in FS\,,\ \sum_{j=1}^{n}(\hat{s}_f(\tau_j) - y_j)^2 \to \min \tag{1}$$

Розв’язок цієї задачі залежить від вхідних даних і може бути записаний як [5, с.140]:

$$\hat{s}_f = P\hat{A}, \tag{2}$$

$$\hat{A} = (P^T P)^{-1} P^T Y\,,$$

де P – матриця планування фрактального сплайна.

Необхідною умовою для існування сплайн-апроксимації є виконання *теореми Шенберга-Уітні*: для кожного $\tau_i \in [a,b]$, $0 \le j < r$ має виконуватись нерівність

$$\xi_j \le \tau_i \le \xi_{j+m}\,, \tag{3}$$

де m - порядок сплайна, ξ - вектор вузлів сплайна. Тобто між вузлами сплайна повинна знаходиться принаймні одна точка даних.

Важливим моментом є *близькість розміщення* вузлів один до одного. Її можна виміряти як

$$\Delta_j = \xi_{j+1} - \xi_j \ \text{ для } 1 \le j \le r-1 \ \text{ і } \ \xi_j \in \xi\,. \tag{4}$$

Величину Δ_j будемо називати ***розміром комірки сітки*** на j -ому інтервалі. Допускається лише така послідовність вузлів, для яких Δ не змінюється дуже швидко, зокрема накладемо умову, щоб

$$\frac{1}{2} \le \frac{\Delta_{j+1}}{\Delta_j} \le 2\,. \tag{5}$$

Також необхідно, щоб розмір комірки на $(k-1)$-ому масштабі був достатнім для заміни фрагмента на k -ому масштабі, тобто

$${}^{k}\Delta \le {}^{k-1}\Delta * \rho\,, \tag{6}$$

де ρ - масштабний множник для k -ого масштабу.

Операцію (2) можна розглядати як часовий низькочастотний фільтр [1, с.2] з шириною каналу

$$BW \cong \frac{1}{2\Delta_j} \text{ для } \xi \text{ біля вузла } \xi_j\,.$$

Оптимальна фільтрація зменшує розмір комірки сітки при збільшенні ширини каналу і збільшує розмір сітки, якщо потрібен вужчий канал.

Алгоритм оптимізації вузлів фрактального сплайна

1. Сформувати початковий набір вузлів.
2. Побудувати базовий сплайн.
3. Побудувати фрактальний сплайн масштабу k .
4. Обчислити помилку регресії для фрактального сплайна.
5. Якщо помилка регресії більше заданого значення, розрахувати нове положення вузлів базового сплайна.
6. Повторити кроки 2-5, доки не буде знайдено мінімальне значення помилки регресії.
7. Побудувати фрактальний сплайн по оптимізованому набору вузлів та обчислити помилку наближення.

Визначення нового положення вузлів сплайна

Пошук оптимальної сітки вузлів здійснюється за допомогою методу покоординатного спуску, який дозволяє враховувати обмеження (5)-(6), вказані вище. Програма циклічно виконує оптимізацію вузлів з другого до передостаннього:

$$\min \Psi(\xi_j, \{\xi_1, ..., \xi_{j-1}, \xi_{j+1}, ..., \xi_r\}), \ j = 2 : r - 1, \tag{7}$$

де $\Psi()$ - цільова функція (сума квадратів відхилень між початковими точками та розрахованими значеннями).

Обмеження області пошуку відрізком Δ_j забезпечує впорядкованість вузлів та наявність даних на кожному відрізку. Точність визначення положення екстремуму обмежена величиною $\Delta / 10$ з метою скорочення обсягів розрахунків.

Приклад роботи алгоритму

Для ілюстрації роботи алгоритму розглянемо оптимізацію вузлів фрактального сплайна на даних валютного курсу євро-долар за період 04.01.1999-19.10.2006, що містить 2000 точок. На рис. 2 (а) представлено фрактальний сплайн масштабу 3, побудований на базисі кубічного ермітового сплайну з 5 вузлами $\xi_1 = [1,500,1000,1500,2000]$. На рис. 2 (б) показано фрактальний сплайн з оптимізованим розміщенням вузлів $\xi_2 = [1,400,1040,1600,2000]$. Середньоквадратична похибка зменшилась на 21,8%, з 0,1032 до 0,0807.

а

б

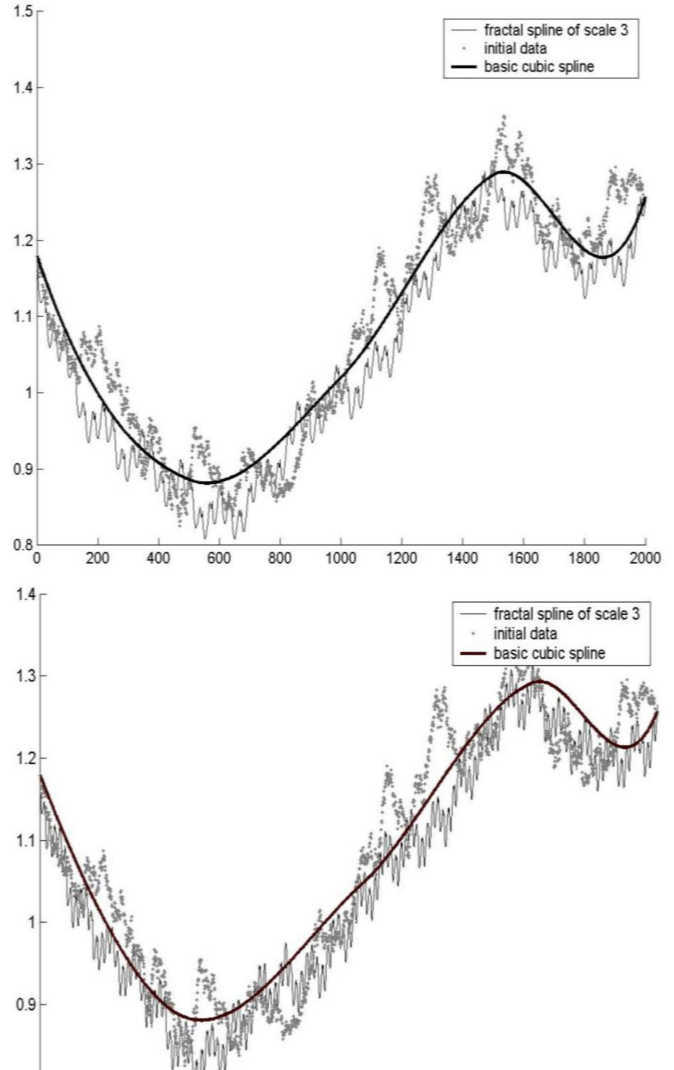

Рис. 2. Оптимізація вузлів фрактального сплайна:
а – апроксимація до оптимізація, б – після оптимізації.

Програмна реалізація алгоритму оптимізаці

Реалізація алгоритму здійснювалась на мові C++. Псевдокод алгоритму наведено у лістингу нижче. На вхід подається початковий набір вузлів (найчастіше рівновіддалених) ξ_temp, вихід — оптимізований набір вузлів ξ_opt. GCV – цільова функція, оцінює похибку регресії за МНК.

```
get ξ_temp
for i=1 to n_search do
 GCV_opt = ∞
 for j=1 to n_sets do
      GCV_temp = ∞
      ξ=ξ_temp + h/10;
      if GCV(ξ) < GCV_temp then
           ξ_temp = ξ
           GCV_temp = GCV(ξ)
      end if
 end for
 ξ_temp := golden section adjustment of ξ_temp
 if GCV(ξ_temp) < GCV_opt then
      ξ_opt = ξ_temp
      GCV_opt = GCV(ξ_temp)
 end if
end for
return ξ_opt
```

Висновки

Отже, існує пряма залежність між вибором вузлів фрактального сплайна та похибкою наближення. *Наукова новизна* даної роботи полягає в тому, що вперше розглянуто питання оптимізації вузлів фрактального сплайна та запропоновано алгоритм вирішення цієї задачі. *Практично значущим результатом* є розробка програми, що виконує оптимізацію в автоматичному режимі. *Предметом подальших досліджень* є удосконалення методу для розрахунку оптимальної кількості вузлів та можливість оцінки віконної апроксимації.

Література

1. Blair J. Optimal knot selection for least-squares fitting of noisy data with spline functions // IEEE Instrumentation and Measurement Technology Conference Proceedings. - Victoria, Vancouver Island, Canada, May 12-15, 2008.
2. Goldenthal R., Bercovier M. Spline Curve Approximation and Design by Optimal Control over the Knots // Computing, Vol. 72, No. 1- 2, pp. 53- 64, April 2004.
3. Spiriti S., Eubank R., Smith P., Young D. Knot Selection for Least Squares and Penalized Splines [Електронний ресурс]. – Режим доступу: http://math.la.asu.edu/~eubank/webpage/vknot.pdf. – Дата доступу: грудень, 2012. – Назва з екрана.
4. Новікова О.Б. Фрактальний сплайн – модель широкосмугового сигналу // Радіоелектроніка та телекомунікації. - №738, 2012. – с. 50-56.
5. Шелевицький І.В. Методи та засоби сплайн-технології обробки сигналів складної форми. – Кривий Ріг: Європейський університет, 2002. – 304 с.