УДК 519.2+681.5.015(0.73.3)

В.В. Турчак, І.В. Шелевицький, В.М. Шутко

«НЕОБМЕЖЕНІ» СПЛАЙНИ В ЗАДАЧАХ ФІЛЬТРАЦІЇ ТА СТИСНЕННЯ ДАНИХ

Розглянуто прості аналітичні вирази одержання оцінок сплайнової моделі методом найменших квадратів за умови, що сітка вузлів сплайна та даних є регулярною і необмеженою в часі.

В системах обробки вимірюваної інформації актуальною є задача стиснення та відновлення даних. При передачі та зберіганні інформації бажано її максимально стиснути. Для фільтрації, аналізу та відображення слід відновити вхідний обсяг даних або навіть розширити його. Один з підходів до розв'язання цієї задачі полягає в передачі та зберіганні лише суттєвих відліків процесу, що вимірюється [1]. При цьому для визначення суттєвості відліків застосовують різні критерії: рівномірний, інтегральний, середньоквадратичний. Відновлення вхідного процесу здійснюється із заданою точністю, тому такі методи відносять до квазіобернених. Фактично алгоритми, які реалізують вказані процедури, формують мінімальну кількість координат параметра, що вимірюється. У випадку відсутності або несуттєвості випадкової складової у вимірюваному процесі задача стиснення зводиться до наближення вхідної послідовності деякою досить простою та зручною в обчисленнях аналітичною функцією. Задача відновлення полягатиме в інтерполяції. Проте досить складно досягти компромісу між точністю наближення та простотою аналітичної функції, особливо для процесів на значних відрізках та складного вигляду. Наявність випадкової складової у вимірюваному процесі, обумовленої як похибками вимірювань, так і внутрішньою природою самого процесу, заставляє застосовувати статистичні методи оцінювання, що ускладнює процедуру стиснення. Тому часто в алгоритмах стиснення використовують алгебраїчні поліноми невисоких порядків у поєднанні з обробкою даних у побіжному вікні або секціями [2].

Привабливий вигляд має застосування із зазначеною метою поліноміальних сплайнів, які при відносній простоті мають хороші апроксимуючі властивості [3]. Проте в загальному випадку застосування сплайнів як аналітичних моделей при статистичному оцінюванні призводить до складних розрахунків, що мало підходять для роботи в реальному часі. Однак правильний вибір сплайна та введення ряду обмежень дозволяють обійти ці труднощі.

Перша проблема полягає у виборі зручного для наближення та розрахунків сплайна. Найбільш відомий – глобальний кубічний сплайн з двома неперервними похідними. Він має властивості найкращої апроксимації та гладкості, проте для побудови навіть для інтерполяції ним потрібно розв'язувати систему лінійних рівнянь. Незважаючи на те, що існують алгоритми побудови такого сплайна з допомогою методу найменших квадратів [4], складність обчислень не дозволяє використати його в реальному часі. Тому найбільш доцільними є локальні сплайни, при розрахунку яких немає потреби розв'язувати систему рівнянь. Локальні інтерполяційні сплайни будують, враховуючи лише найближчі вузли, і вони не мають оптимальних властивостей глобальних сплайнів. Кубічні ермітові сплайни мають неперервну першу похідну, визначену з умови відповідності її значення реальному процесові. Це зумовлює їхню гладкість і хорошу відповідність процесові, який наближають. В лагранжевій формі запису кубічні інтерполяційні сплайни просто визначаються через вузли – точки стикування фрагментів.

Друга проблема – необмеженість вхідного потоку даних. У самому визначенні сплайн вважається функцією, заданою на певному обмеженому інтервалі. Для глобальних сплайнів подолати це обмеження неможливо. Локальні, якщо не враховувати незначну початкову похибку (перехідний процес), можна будувати без обмежень на обсяг даних.

Отже, розглянемо алгоритм оцінювання сплайн-моделі методом найменших квадратів за даними вимірювань. Якщо сітка вузлів сплайна регулярна і на кожному з фрагментів сплайна N даних розміщені однаковим чином : $\beta_i = t_j - x_i$, $i = \overline{1,N}$, то такий сплайн з необмеженим числом вузлів $R \to \infty$ повністю описується одним фрагментом. Для ермітового кубічного сплайна значення в довільній i-й точці регулярної сітки даних

$$S(x_i) = A_{j-1}\varphi_{j-1}(x_i) + A_j\varphi_j(x_i) + A_{j+1}\varphi_{j+1}(x_i) + A_{j+2}\varphi_{j+2}(x_i),$$

де A_j – значення сплайна в j-у вузлі ; φ_i – локальна функція форми.

Функція форми має деякі властивості.

1. Функція форми є сплайном. Вона складається з чотирьох фрагментів, які є поліномами третього ступеня і узгоджені в точках стикування власними значеннями та значеннями першої похідної.

2. Функція форми має ненульові значення лише на чотирьох сусідніх фрагментах.

3. Сума всіх функцій форми в довільній точці, що належить сплайну $t \in [t_0, t_R]$, дорівнює одиниці.

4. Функції форми інваріантні до одночасного перетворення абсцис tu та t відповідно до виразу

$$\begin{cases} t_i = t_i + h; \\ tu_j = tu_j + h, \end{cases} \text{де } h \in (-\infty, \infty), \ i = \overline{1,N}, \ j = \overline{0,R}.$$

5. Функції форми інваріантні до одночасного перетворення абсцис tu та t відповідно до виразу

$$\begin{cases} t_i = mt_i; \\ tu_j = mtu_j, \end{cases} \text{де } m \in (0, \infty), \ i = \overline{1,N}, \ j = \overline{0,R}.$$

6. Функція форми φ_i приймає максимальне значення, рівне одиниці, в точці tu_j.

7. Функція форми φ_i приймає нульові значення всередині ненульового інтервалу в точках $tu_{j-2}, tu_{j-1}, tu_{j+1}$.

8. Функція форми φ_i симетрична відносно точки tu_j.

Вигляд функції форми показаний на рис.1.

Врахувавши характер функції форми, запишемо:

$$S_i = A_{j-1}g1_i + A_jg2_i + A_{j+1}g3_i + A_{j+2}g4_i, \tag{1}$$

де

$$g1_i = -\frac{1}{2N^3}\left(i^3 - 2Ni^2 + N^2 i\right);$$

$$g2_i = \frac{1}{2N^3}\left(3i^3 - 5Ni^2 + 2N^3\right);$$

$$g3_i = \frac{1}{2N^3}\left(-3i^3 + 4Ni^2 + N^2 t\right);$$

$$g4_i = -\frac{1}{2N^3}\left(-i^3 + Ni^2\right) i = \overline{1,N}.$$

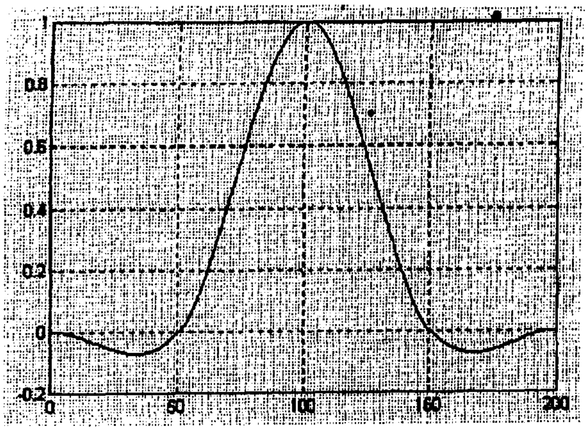

Рис. 1

Тепер можна записати систему з N лінійних рівнянь $S=GA$:

$$\begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_N \end{bmatrix} = \begin{bmatrix} g1_1 & g2_1 & g3_1 & g4_1 \\ g1_2 & g2_2 & g3_2 & g4_2 \\ \dots & \dots & \dots & \dots \\ g1_N & g2_N & g3_N & g4_N \end{bmatrix} \begin{bmatrix} A_{j-1} \\ A_j \\ A_{j+1} \\ A_{j+2} \end{bmatrix}$$

Для необмеженого числа фрагментів матриця G буде блочно-діагональною і складатиметься з однакових ненульових блоків для фрагментів. Щоб отримати оцінки сплайна у вузлах за методом найменших квадратів за необмеженою кількістю даних F запишемо систему нормальних рівнянь:

$$G^*GA=G^*F$$

або позначивши інакше,

$$CA=B.$$

Матриця C – симетрична семидіагональна матриця, яка повністю визначається чотирма ненульовими елементами довільного рядка

$$c_{j,j} = c_0, c_{j,j+1} = c_{j+1,j} = c_1,\ c_{j,j+2} = c_{j+2,j} = c_2,\ c_{j,j+3} = c_{j+3,j} = c_3;$$

$$c_0 = \frac{171N^6 + 14N^2 + 25}{210N^5};$$

$$c_1 = \frac{71N^6 - 21N^2 - 50}{560N^5};$$

$$c_2 = -\frac{N^6 - 1}{28N^5};$$

$$c_3 = \frac{3N^6 + 7N^2 - 10}{1680N^5}.$$

За таких умов матриця C^{-1} є симетричною з діагональним переважанням. Можемо вважати, що $C^{-1} \approx D$, де остання є семидіагональною симетричною матрицею з елементами, рівними (d_0- діагональний елемент) :

$$d_0 = \left(c_0^3 + c_0^2 c_2 - 2c_0 c_1^2 - 2c_0 c_1 c_3 - c_0 c_2^2 - c_0 c_3^2 + c_1^2 c_2 + 2c_1 c_2 c_3\right)/Q;$$
$$d_1 = -\left(c_0^2 c_1 - c_0 c_1 c_2 - 2c_0 c_2 c_3 - c_1^3 + c_1^2 c_3 + c_1 c_2^2 + c_1 c_3^2\right)/Q;$$
$$d_2 = -\left(c_0^2 c_2 - c_0 c_1^2 - 2c_0 c_1 c_3 - c_2^3 + c_0 c_2^2 + c_1^2 c_2 + c_2 c_3^2\right)/Q;$$
$$d_3 = -\left(c_0^2 c_3 - 2c_0 c_1 c_2 + c_0 c_2 c_3 + c_1^3 - 2c_1 c_3^2 + c_2^2 c_3 - c_3^3\right)/Q,$$

де

$$Q = c_0^4 + c_0^3 c_2 - 4c_0^2 c_1^2 - 2c_0^2 c_1 c_3 - 3c_0^2 c_2^2 - 3c_0^2 c_3^2 + 5c_0 c_1^2 c_2 + 14 c_0 c_1 c_2 c_3 - 2c_0 c_2^3 -$$
$$- 2c_0 c_2 c_3^2 + 2c_1^4 - 4c_1^3 c_3 - 4c_1^2 c_2^2 - 2c_1^2 c_3^2 + 4c_1 c_3^3 + 2c_2^4 - 4c_2^2 c_3^2 + 2c_3^4.$$

Оцінка значення сплайна в довільному вузлі розраховується за формулою

$$a_j = d_0 b_j + \sum_{i=1}^{3} d_i \left(b_{j-1} + b_{j+1}\right).$$

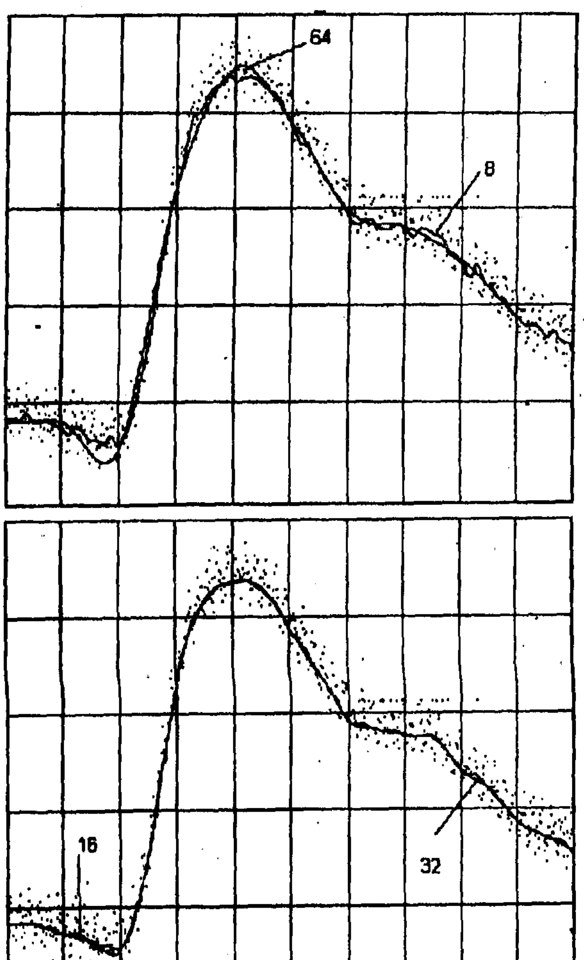

Рис. 2

Таким чином, описаний метод дозволяє стискати вхідні дані в N разів. Відновлення вхідного числа даних здійснюється шляхом інтерполяції за отриманими оцінками у відповідності з виразом (1). Простота обчислень дозволяє реалізувати метод і апаратурно. Можливий ступінь стиснення залежить від диференційних властивостей корисного сигналу, вимог до точності наближення та якості статистичних оцінок. Розширення фрагментів сплайна із збільшенням N поліпшує статистичні оцінки (зменшується d_0), але збільшує похибку наближення сплайном детермінованої складової вхідного процесу у відповідності з виразом $\frac{\max\left|f^{(4)}(t)\right|}{384} h^4$, де f – детермінована основа вхідного процесу. Тому потрібно знаходити компромісне значення.

Як приклад розглянемо стиснення даних в медичній функціональній діагностиці. Система обробки отримує електрокардіографічні та реографічні дані з аналогового пристрою і оцифровує їх з частотою 1033 Гц. Частота вибрана виходячи з частотного діапазону аналогового каналу та особливостей системи відображення даних. Корисна інформаційна складова реосигналу має значно вужчий частотний діапазон, ніж аналоговий канал. Виходячи з цього, можна очікувати значного стиснення даних з реалізацією викладеного методу. На рис.2 показано вигляд апроксимуючої сплайн-функції при стисненні в 8, 16, 32 та 64 рази.

При стисненні у вісім разів гладкість сплайна недостатня. Модель повторює випадковий процес. Стиснення в 64 рази дає надто згладжений результат. Оптимальним є стиснення в 16–32 рази. Враховуючи паралельну обробку кардіосигналу, в системі використано стиснення в 16 разів. Зауважимо, що просте аналітичне представлення вхідного сигналу значно спростило його подальший аналіз (диференціювання та пошук екстремумів).

Список літератури

1. *Дядюнов А.Н., Онищенко Ю.А., Сенин А.И.* Адаптивные системы сбора и передачи аналоговой информации. – М.: Машиностроение, 1988. – 253 с.

2. *Цифровая* обработка сигналов /Гольденберг Л.М. и др. – М.: Радио и связь, 1985. – с.382.

3. *Корнейчук Н.П.* Сплайны в теории приближения. – М.: Наука, 1984. – 211 с.

Стаття надійшла в редакцію 5 лютого 1998 року.

Віталій Володимирович Турчак (1969) закінчив Київський інститут інженерів цивільної авіації в 1993 році. Асистент кафедри авіаційних радіоелектронних систем Київського міжнародного університету цивільної авіації. Основний напрямок наукових досліджень – розробка та дослідження сплайн-методів обробки інформації. Автор чотирьох статей.

Vitalyi V. Turchak (b.1969) graduated from Kyiv Institute of Civil Aviation Engineers (1993). Instructor of radio electronic aviation systems department of Kyiv International University of Civil Aviation, the fields of inter[illegible] [illegible] information processing. Author of four publications.

Ігор Володимирович Шелевицький (1963) закінчив Київський інститут інженерів цивільної авіації в 1986 році. Кандидат технічних наук, докторант Київського міжнародного університету цивільної авіації. Основний напрямок наукових досліджень – розробка та дослідження сплайн-методів обробки інформації. Автор дев'яти статей.

Igor V. Shelevitsky(b. 1963) graduated from Kyiv Institute of Civil Aviation Engineers (1986). PhD (Eng), doctorate of Kyiv International University of Civil Aviation. Specializes in the field of spline-methods of information processing. Author of 9 publications.

Володимир М[illegible] Шутко (1970) закінчив Московський державний технічний університет ім. М.Е. Баумана в 1993 році. Асистент кафедри радіоелектроніки Київського міжнародного університету цивільної авіації. Основний напрямок наукових д[illegible] – р[illegible] [illegible]айн-методів обробки інформації. Автор вісьми статей.

Volodymir M. Shutko (b. 1970) graduated from Bauman Moscow Higher Technical University in 1993, instructor of radio electronics departmens of Kyiv International University of Civil Aviation, the fields of interest are: wo[illegible] [illegible] information. Author of 8 publications.