УДК 517.9

І. В. Абрамчук

РЕГРЕСІЯ НА ОСНОВІ SINC-ФУНКЦІЇ

Вінницький національний технічний університет

Анотація.
Запропоновано метод обробки експериментальних даних на основі методу найменших квадратів, де апроксимуюча функція будується як лінійна комбінація модифікованих sinc-функцій. Матриця системи нормальних рівнянь запропонованого алгоритму має меншу обумовленість ніж, наприклад, у випадку поліноміальної регресії, це дозволяє використовувати більшу кількість доданків апроксимуючої функції, без катастрофічної втрати точності обчислень.
Ключові слова: метод найменших квадратів, регресія, матриця системи нормальних рівнянь, sinc-функція, погано обумовлена СЛАР.

Abstract.
A method of processing experimental data based on the least square's method is proposed, where the approximating function is constructed as a linear combination of modified sinc functions. The matrix of normal equations system of the proposed algorithm has less condition than, for example, in the polynomial regression case, this allows the use of more additives in the approximating function, without catastrophic loss of calculation accuracy.
Keywords: least squares method, regression, the matrix of normal equations system, sinc function, poorly conditioned SLAE.

Для знаходження аналітичної залежності між величинами за результатами серії експериментів, особливо у випадку коли експериментальні дані отримані з деякими похибками, що не мають систематичного характеру, використовують апроксимуючі функції, що мінімізують певну норму похибки. Одним з найрозповсюдженіших алгоритмів розв'язування таких задач є метод найменших квадратів [1, 2]. Якщо у якості апроксимуючої функції вибирати лінійну комбінацію базисних функцій, то задача зводиться до розв'язування СЛАР [1, 2].

Постановка задачі

Для заданої системи базисних функцій $\{\varphi_k(x)\}_{k=0..(m-1)}$ і вектора експериментальних даних $\{(x_i, y_i)\}_{i=0..(n-1)}$, де $m \ll n$ визначити вектор $\vec{c}$, що мінімізує функціонал

$$\Delta_2(x,\vec{c}) = \sum_{i=0}^{n} \left(\Phi(x_i,\vec{c}) - y_i\right)^2 \to \min_{\vec{c}}, \text{ де } \Phi(x,\vec{c}) = \sum_{k=0}^{m} c_k \varphi_k(x), \quad (1)$$

Запропонувати вибір базисних функцій $\varphi_k(x)$, так, щоб матриця $M^T M$ системи нормальних рівнянь [1,2]

$$M^T M \cdot \vec{c} = M^T \vec{y}, \quad (2)$$

де $M_{i,k} = \varphi_k(x_i)$, $\vec{y} = \{y_i\}_{i=0..(n-1)}$, для визначення оптимальних параметрів $\vec{c}$ задачі (1), була б якомога *добре обумовленою*.

Розв'язання поставленої задачі

Метод (1) є лінійною формою метода найменших квадратів (МНК) апроксимації вектора даних.

Виберемо у якості базисних функцій $\varphi_k(x) = \operatorname{sinc}\left((1/\Delta x)\pi(x - k\cdot\Delta x)\right)$, де $\operatorname{sinc}(x) = \sin x / x$ при $x \neq 0$ і $\operatorname{sinc}(x) = 1$ при $x = 0$, Δx – крок для базисних функцій, вибраний так, що $k \cdot \Delta x$ не співпадає з абсцисами вектора даних. Функції $\varphi_k(x)$ володіють властивостями подібними до локальних властивостей В- сплайнів: $\varphi_k(x = k\cdot\Delta x) = 1$ і $\varphi_k(x = i\cdot\Delta x) = 0$, $i \neq k$, вони всюди диференційовані.

Такого роду функції успішно використовуються при розв'язанні задач інтерполяції, чисельного розв'язування задач з початковими і крайовими умовами для звичайних диференціальних рівнянь, розв'язуванні інтегральних рівнянь, тощо [3,4,5].

Коефіцієнти $\vec{c}$ методу найменших квадратів (1) знаходяться з системи нормальних рівнянь (2), яку можна розв'язувати, наприклад, подаючи симетричну, додатно визначену матрицю $M^T M$ у вигляді QR або SVD факторізації. Відомо [1,2], що коли в якості базисних функцій вибрано степеневі (поліноміальна регресія), то обумовленість матриці $M^T M$ росте надзвичайно швидко і при $m > 16$ похибки машинних обчислень практично унеможливлюють знаходження розв'язку задачі оптимізації [1,2]. Якщо поліноміальний базис замінити на базис на основі sinc – функцій, то, як показують експерименти, обумовленість матриці $M^T M$ суттєво зменшується, що цілком відповідає характеру поведінки і властивостям цих функцій [3-5]. Розглянемо приклад

Приклад. Нехай задана функція $f(x) = (\sin x + 3\cos 2x) \cdot e^{-0.5x}$, яка в точках $x_i = 0 + i \cdot \delta x$, $i = 0..(n-1)$, $n = 2^7$ спотворена гаусовим шумом з нульовим математичним сподіванням і заданим СКВ σ. Будемо шукати коефіцієнти апроксимуючої функції $\Phi(x, \vec{c})$ з m параметрами, $m = 2^4$ (їх кількість визначаємо за принципом Найквіста). Якщо у якості базисних функцій взяти sinc-функції, що відповідають точкам $x_k = 0 + k \cdot \Delta x$, $k = 0..(m-1)$, то число зумовленості матриці $\operatorname{cond}\left(M^T M\right) \approx 1.9 \cdot 10^1$, при $\sigma \in [0;3]$, а відхилення апроксимуючої функції від заданої є цілком допустимим. Випадок, коли вибрано $\sigma = 0.5$, ілюструє рисунок.

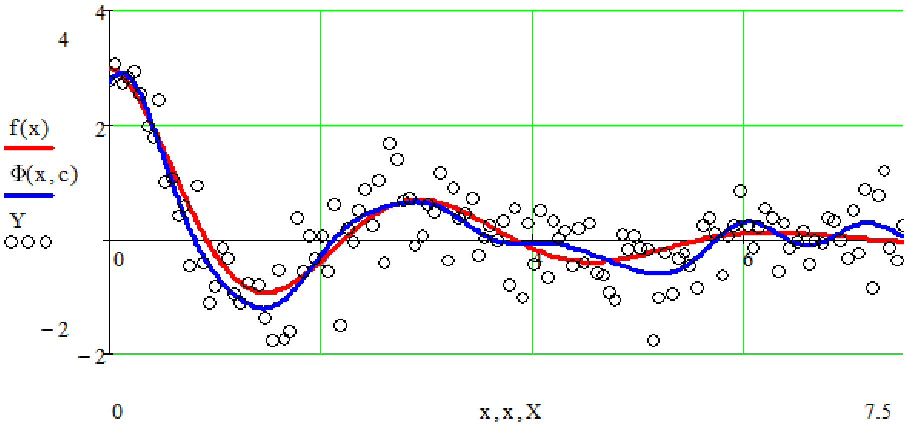

Рис. Результат МНК апроксимації на основі sinc-функцій при $m = 2^4$, $\sigma = 0.5$.

У випадку використання поліноміальної регресії $\operatorname{cond}\left(M^T M\right) \approx 1.0 \cdot 10^{18}$ і апроксимуюча функція навіть віддалено не нагадує вхідну неспотворену функцію $f(x)$.

Висновки

Запропоновано метод обробки експериментальних даних на основі методу найменших квадратів, де апроксимуюча функція будується як лінійна комбінація модифікованих sinc-функцій. Матриця системи нормальних рівнянь запропонованого алгоритму має меншу обумовленість ніж, наприклад, у випадку поліноміальної регресії, а це дозволяє використовувати більшу кількість доданків апроксимуючої функції, без катастрофічної втрати точності обчислень.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Райс Дж. Матричные вычисления и математическое обеспечение / Дж. Райс; пер. с англ. – М.: Мир, 1984. – 264 с.
2. Деммель Дж. Вычислительная линейная алгебра / Дж. Деммель; пер. с англ. – М.: Мир, 2001. – 429 с.

3. Sugihara M., Matsuo T. Recent developments of the Sinc numerical methods, // J. Comput. Appl. Math. 164–165 (2004) 673 – 689.
4. Stenger F. Numerical Methods Based on Sinc and Analytic Functions, / Springer Series in Computational Mathematics, Vol. 20, Springer-Verlag, New York, 1993.
5. Stenger F. Summary of Sinc numerical methods // J. Comput. Appl. Math. 121 (2000) 379-420

Абрамчук Ігор Васильович, старший викладач кафедри вищої математики, Вінницький національний технічний університет, м. Вінниця, email: igrabramchuk@gmail.com

Abramchuk Igor V., department of higher mathematic, Vinnitsa National Technical University, Vinnitsa, email: igrabramchuk@gmail.com