ISSN 0002-306X. Изв. НАН РА и НПУА. Сер. ТН. 2015. Т. LXVIII, N2.

УДК 621.52+519.6

АВТОМАТИЗАЦИЯ И СИСТЕМЫ УПРАВЛЕНИЯ

Ю.Р. АКОПЯН, М.Г. ХАЧАТРЯН

ОБ ИСПОЛЬЗОВАНИИ КУБИЧЕСКИХ СПЛАЙНОВ В ЗАДАЧЕ СГЛАЖИВАНИЯ ДАННЫХ

Рассматривается задача сглаживания экспериментальных данных методом наименьших квадратов с использованием кубических сплайнов с краевыми условиями смешанного типа. Предлагается численный алгоритм построения таких сплайнов. Дается описание созданного в среде Matlab пользовательского интерфейса.

Ключевые слова: сглаживание данных, метод наименьших квадратов, нормальная система, кубический сплайн, пакет MatLab, графический интерфейс.

Введение. Современный уровень развития компьютерной техники позволяет решать практические задачи, связанные с обработкой больших массивов данных. Такие задачи возникают во многих областях, использующих математические методы. В частности, хорошо известна задача сглаживания результатов статистических наблюдений, экспериментов или измерений. Несмотря на свою давность, метод наименьших квадратов остается одним из наиболее эффективных и широко используемых на практике методов математической обработки опытных данных. Многие пакеты прикладных программ, например, такие как MatLab, Mathematica и Mathcad, включают в себя инструменты для решения задачи наименьших квадратов. Однако следует отметить, что набор аппроксимирующих функций и предоставляемые пользователю возможности еще далеки от совершенства. В частности, это касается выбора близкой к характеру поведения данных аппроксимирующей функции, графического представления результатов сглаживания, возможности оперативного изменения параметров и т.д.

Как известно, сплайны являются важной составной частью математической теории аппроксимации и имеют многочисленные применения в различных вычислительных приложениях [1,2]. Наиболее широко используются на практике, в силу простоты их построения и достаточно высокой гладкости, полиномиальные кубические сплайны. Вместе с тем они недостаточно полно представлены в программных средствах, ориентированных на решение задачи сглаживания данных методом наименьших квадратов.

В существующей литературе рассматриваются кубические сплайны с различными ограничениями на концах отрезка аппроксимации. Обычно это условия, накладываемые на первые и вторые производные сплайнфункции. В настоящей работе рассматриваются кубические сплайны с более общими, смешанными краевыми условиями, содержащими комбинацию первых и вторых производных. Приводится алгоритм построения таких сплайнов. Дается описание работающего в диалоговом режиме пользовательского интерфейса в среде MatLab, позволяющего сглаживать данные методом наименьших квадратов на основе кубических сплайнов с краевыми условиями смешанного типа.

Кубические сплайны с краевыми условиями смешанного типа. Рассмотрим некоторое разбиение

$$a = t_0 < t_1 < \cdots < t_n = b \qquad (1)$$

отрезка $[a,b]$, где $n \geq 1$. Пусть в точках разбиения t_k, $\mathrm{k} = 0,1,\ldots,\mathrm{n}$ заданы значения v_k, $k = 0,1,\ldots,n$. Функция $S(x)$ называется *интерполяционным кубическим сплайном,* ассоциированным с разбиением (1), если:

а) на каждом отрезке $[t_k, t_{k+1}]$, $k = 0,1,\ldots,n-1$ функция $S(x)$ является многочленом третьей степени;

б) $S(x) \in C^2[a,b]$;

в) $S(t_k) = v_k$, $k = 0,1,\ldots,n$

(см., напр., [1]). Задаваемые условия обеспечивают непрерывность самой функции $S(x)$, а также ее первой и второй производных в точках разбиения (1), которые в дальнейшем будем называть *опорными узлами* сплайна.

Как известно, для корректности задачи построения сплайна нужны дополнительные условия [2]. Обычно эти условия ставятся на концах отрезка $[a,b]$ и содержат требования на первые и вторые производные функции $S(x)$. Если $S'(a) = d_a$, $S'(b) = d_b$, где $\mathrm{d_a}, \mathrm{d_b}$ – заданные величины, то такой сплайн называется *смыкающим.* Условия $S''(a) = d_a$, $S''(b) = d_b$ определяют кубический сплайн, имеющий заданную кривизну на концах отрезка. В частности, если $S''(a) = 0$, $S''(b) = 0$, то такой сплайн называется *естественным* (см., напр., [3]). В настоящей работе мы зададим более общий, *смешанный тип* краевых условий, а именно:

$$\alpha_1 S''(a) + \alpha_2 S'(a) = d_a, \quad \beta_1 S''(b) + \beta_2 S'(b) = d_b; \qquad (2)$$

при этом мы полагаем, что коэффициенты α_1, α_2 и β_1, β_2 таковы, что

$$\alpha_1 \geq 0,\ \alpha_2 \leq 0,\ \alpha_1^2 + \alpha_2^2 \neq 0,\ \beta_1 \geq 0,\ \beta_2 \geq 0,\ \beta_1^2 + \beta_2^2 \neq 0. \qquad (3)$$

В работе [4] получено аналитическое выражение для кубического сплайна с краевыми условиями (2). Для значений $k = 0,1,\ldots,n-1$ сплайн вычисляется по формуле

$$S(x) = v_k + \left(\frac{v_{k+1} - v_k}{h_k} - \frac{2M_k + M_{k+1}}{6} h_k \right)(x - t_k) + \\ + \frac{M_k}{2}(x - t_k)^2 + \frac{M_{k+1} - M_k}{6h_k}(x - t_k)^3, \quad x \in [t_k, t_{k+1}], \tag{4}$$

где $h_k \equiv t_{k+1} - t_k$, $M_k \equiv S_k''(t_k)$, $k = 0,1,\ldots,n-1$; $M_n \equiv S_{n-1}''(t_n)$.

Величины M_k определяются из трехдиагональной системы линейных алгебраических уравнений

$$\begin{cases} d_1 M_1 + b_1 M_2 = g_1, \\ a_k M_{k-1} + d_k M_k + b_k M_{k+1} = g_k, \quad k = 2,3,\ldots,n-2, \\ a_{n-1} M_{n-2} + d_{n-1} M_{n-1} = g_{n-1}, \end{cases} \tag{5}$$

коэффициенты которой вычисляются по формулам

$$d_1 = 2h_1 + \frac{3h_0(4\alpha_1 - \alpha_2 h_0)}{2(3\alpha_1 - \alpha_2 h_0)}, \quad b_1 = h_1,$$

$$a_k = h_{k-1}, \quad d_k = 2(h_{k-1} + h_k), \; b_k = h_k, \quad k = 2,3,\ldots,n-2,$$

$$a_{n-1} = h_{n-2}, \quad d_{n-1} = 2h_{n-2} + \frac{3h_{n-1}(4\beta_1 + \beta_2 h_{n-1})}{2(3\beta_1 + \beta_2 h_{n-1})},$$

а правые части – по формулам

$$g_1 = 6\left(\frac{v_2 - v_1}{h_1} - \frac{v_1 - v_0}{h_0} \right) - \frac{3h_0}{3\alpha_1 - \alpha_2 h_0}\left(d_a - \alpha_2 \frac{v_1 - v_0}{h_0} \right),$$

$$g_k = 6\left(\frac{v_{k+1} - v_k}{h_k} - \frac{v_k - v_{k-1}}{h_{k-1}} \right), \; k = 2,3,\ldots,n-2,$$

$$g_{n-1} = 6\left(\frac{v_n - v_{n-1}}{h_{n-1}} - \frac{v_{n-1} - v_{n-2}}{h_{n-2}} \right) - \frac{3h_{n-1}}{3\beta_1 + \beta_2 h_{n-1}}\left(d_b - \beta_2 \frac{v_n - v_{n-1}}{h_{n-1}} \right).$$

Система уравнений (5) эффективно решается с помощью известного метода прогонки [5]. При этом особо подчеркнем, что при наличии условий (3) матрица системы (5) имеет строгое диагональное преобладание, а именно:

$$|d_1| > |b_1|,$$
$$|d_k| > |a_k| + |b_k|, \; 2 \le k \le n-2,$$
$$|d_{n-1}| > |a_{n-1}|.$$

Это обеспечивает корректность и устойчивость вычислительного алгоритма метода прогонки [5].

Сглаживание данных методом наименьших квадратов. Пусть зависимость между двумя переменными величинами x и y задается в виде таблицы

$$\frac{x}{y}\left\|\frac{x_0}{y_0}\right|\frac{x_1}{y_1}\left|\frac{\cdots}{\cdots}\right|\frac{x_m}{y_m}. \tag{6}$$

Выберем отрезок аппроксимации $[a,b]$ так, чтобы $a \le \min\limits_{0\le i\le m} x_i$, $b \ge \max\limits_{0\le i\le m} x_i$. Зададим некоторое разбиение отрезка $[a,b]$ типа (1). Сформулируем следующую задачу: *построить ассоциированный с разбиением (1) кубический сплайн* $S(x)$ *с краевыми условиями смешанного типа (2), для которого сумма квадратов отклонений*

$$E(v_0, v_1, \ldots, v_n) \equiv \sum_{i=0}^{m} \omega_i \left[S(x_i) - y_i\right]^2 , \tag{7}$$

где ω_i *есть заданный вес в узле* t_i *, принимает наименьшее значение.*

Кубический интерполяционный сплайн $S(x)$ определяется своими значениями $v_0, v_1, \ldots, v_n$ в опорных узлах разбиения (1). Тем самым минимизируемая величина зависит от этих значений. Следовательно, поставленную задачу (7) можно переформулировать следующим образом: *найти величины* $v_0^*, v_1^*, \ldots, v_n^*$*, при которых*

$$E(v_0^*, v_1^*, \ldots, v_n^*) = \min_{v_0, v_1, \ldots, v_n} E(v_0, v_1, \ldots, v_n). \tag{8}$$

Построение численного алгоритма решения задачи (8) и его теоретическое обоснование даны в работе [6]. Ниже приводится краткое описание основных этапов алгоритма.

Этап 1. Выявление зависимости величин M_k , $k = 0,1,\ldots,n$, определяемых из системы (5), от величин $v_0, v_1, \ldots, v_n$.

Получены выражения

$$M_k = \sum_{j=0}^{n} \gamma_{kj} v_j + \sigma_k , \quad k = 0,1,...,n$$

для величин M_k , где коэффициенты γ_{kj} и σ_k вычисляются по формулам, зависящим от параметров краевых условий (2) и элементов обратной матрицы H^{-1} , где H есть трехдиагональная матрица системы (5). Элементы матрицы H^{-1} вычисляются по выведенным явным формулам.

Этап 2. Выявление зависимости значений сплайна $S(x_i)$, $i = 0,1,...,m$ от величин $v_0, v_1, \ldots, v_n$.

Получены выражения

$$S(x_i) = \sum_{j=0}^{n} a_{ij} v_j + \rho_i, \qquad i = 0,1,...,m,$$

где коэффициенты a_{ij} и ρ_i вычисляются через величины, полученные на предыдущем Этапе 1, а именно, через величины γ_{kj} и σ_k.

Этап 3. Вычисление оптимальных значений величин $v_0, v_1,..., v_n$ (решение задачи (8)).

Имея величины, вычисленные на Этапах 1 и 2, задача построения сплайна сводится к решению *нормальной* системы

$$A^T \Omega A v = A^T \Omega (y - \rho), \tag{9}$$

где $A = [a_{ij}]_{i=0,1,...m}^{j=0,1,...n}$, $\Omega = diag[\omega_0, \omega_1,...,\omega_m]$ и $\mathrm{y} = [\mathrm{y}_0, \mathrm{y}_1,..., \mathrm{y}_\mathrm{m}]^\mathrm{T}$, $\rho = [\rho_0, \rho_1,..., \rho_\mathrm{m}]^\mathrm{T}$, $\mathrm{v} = [\mathrm{v}_0, \mathrm{v}_1,..., \mathrm{v}_\mathrm{n}]^\mathrm{T}$.

В работе [6] были исследованы свойства матрицы системы (9). А именно, если $m > 3n$ и разбиение (1) отрезка $[a,b]$ таково, что каждый интервал (t_k, t_{k+1}), $0 \le k \le n-1$ содержит по крайней мере три точки x_i из таблицы данных (6), то матрица $A^T \Omega A$ является положительно - определенной.

Описание программного пакета и пользовательского интерфейса. На основе полученных теоретических результатов в среде MatLab был создан программный пакет и работающий в диалоговом режиме графический интерфейс GUI (Graphical User Interface). При этом были использованы общие принципы и методы создания подобных интерфейсов [7,8]. Основное назначение программного пакета – решение задачи сглаживания данных с использованием не только кубических сплайнов, но и алгебраических многочленов, показательных (экспоненциальных), логарифмических и дробно-рациональных функций.

Графический интерфейс GUI управляется с помощью созданных утилитов, в том числе кнопок, меню и других инструментов, представленных в графическом окне (см. рис.). Там же представлены вводимые эсперимиентальные данные и полученные в результате работы вычислительных алгоритмов аппроксимирующие кривые.

После запуска программы перед пользователем открывается окно графического интерфейса, в котором необходимо либо ввести сглаживаемые данные, либо вызвать файл, с которого программа их считывает. Затем, если рассматривается аппроксимация кубическими сплайнами, вводятся начальные значения параметров краевых условий смешанного типа. Если пользователь в качестве начального данного введет недопустимое значение, то откроется диалоговое окно, в котором появляется предупреждение об ошибке. При нажатии соответствующей кнопки в графическом окне появляется распределение данных. С помощью последовательных нажатий мышки пользователь может зафиксировать опорные узлы кубического сплайна. В соответствующем поле окна появляется число уже отмеченных опорных узлов; в других полях видны текущие координаты курсора.

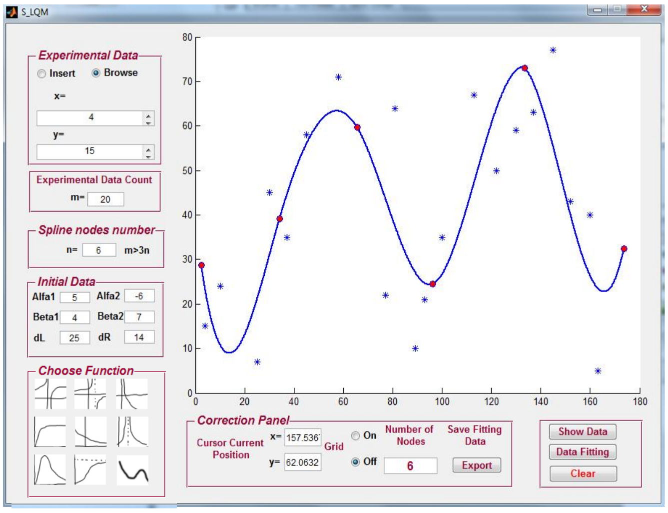

Рис. Аппроксимирующая кривая в окне пользовательского интерфейса

После того как опорные узлы выбраны, запускается программа, и в графическом окне появляется сглаживающая кривая (рис.). Одна из главных особенностей работы интерфейса заключается в том, что пользователь имеет возможность в интерактивном режиме с помощью мышки передвигать опорные узлы кубического сплайна. С изменением опорных узлов соответственно меняется аппроксимирующая кривая. Предусмотрена также кнопка интерфейса, которая экспортирует некоторые данные в файл (например, величины отклонений и координаты опорных узлов).

Заключение. Рассмотренные в работе кубические сплайны с краевыми условиями смешанного типа и полученные результаты могут послужить поводом для более детального теоретического изучения таких функций. Кубические сплайны, включенные в разработанный графический интерфейс, могут использоваться при статистических исследованиях и решении инженерных задач.

СПИСОК ЛИТЕРАТУРЫ

1. **Алберг Дж., Нильсон Э., Уолш Дж.** Теория сплайнов и ее приложения.– М.: Мир, 1972. – 318 с.
2. **Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л.** Методы сплайн-функций.– М.: Наука, 1980. – 352 с.
3. **Мэтьюз Дж., Финк К.Д.** Численные методы: использование MATLAB.– М.: Издательский дом “Вильямс”, 2001. – 720 с.
4. **Խաչատրյան Մ.Գ.** Խառը տիպի եզրային սահմանափակումներով խորանարդային սպլայնի կառուցումը // ՀՊՃՀ Բանբեր. Մոդելավորում, օպտիմալացում, կառավարում. – 2011. – Թող. 14, հ.2. – էջ 33-40:
5. **Самарский А.А., Николаев Е.С.** Методы решения сеточных уравнений.– М.: Наука, 1978. – 592 с.
6. **Khachatryan M.** Least Squares Fitting with Cubic Splines // Mathematical Problems of Computer Science, Transactions of the Institute for Informatics and Automation Problems of the National Academy of Sciences of Republic of Armenia. – 2012. – V.37. – P. 53-53.
7. **Scott T. S.** MATLAB: Advanced GUI Development. – Dog Ear Publ., Indianapolis, 2006. – 304 p.
8. **Ермачкова Ю.А.** Проектирование интерфейса в среде GUIDE MATLAB // Современные информационные технологии в экономике, управлении и образовании: Сборник материалов Межвузовской научно-практической конференции.– М., 2006. – С. 35-37.

Ереванский государственный университет, Национальный политехнический университет Армении. Материал поступил в редакцию 03.02.2015.

ՅՈՒ.Ռ. ՀԱԿՈԲՅԱՆ, Մ.Գ. ԽԱՉԱՏՐՅԱՆ

ՏՎՅԱԼՆԵՐԻ ՀԱՐԹԵՑՄԱՆ ԽՆԴՐՈՒՄ ԽՈՐԱՆԱՐԴԱՅԻՆ ՍՊԼԱՅՆՆԵՐԻ ԿԻՐԱՌՈՒԹՅԱՆ ՄԱՍԻՆ

Փոքրագույն քառակուսիների մեթոդով դիտարկվում է փորձարարական տվյալների հարթեցման խնդիրը՝ խառը տիպի եզրային պայմաններով խորանարդային սպլայնների կիրառմամբ: Առաջարկվում է այդպիսի սպլայնների կառուցման թվային ալգորիթմ: Տրվում է Matlab միջավայրում մշակված օգտագործողի ինտերֆեյսի նկարագրությունը:

Առանցքային բառեր. տվյալերի հարթեցում, փոքրագույն քառակուսիների մեթոդ, նորմալ համակարգ, խորանարդային սպլայն, MatLab փաթեթ, գրաֆիկական ինտերֆեյս:

Yu.R. HAKOPIAN, M.G. KHACHATRYAN

THE USE OF CUBIC SPLINES IN THE DATA FITTING PROBLEM

The experimental data fitting problem by the least squares method using cubic splines with mixed type boundary conditions is considered. A numerical algorithm for constructing such splines is proposed. A user interface created in Matlab environment is introduced.

Keywords: data fitting, least squares method, normal system, cubic spline, MatLab system, graphical interface.