

НАПРАВЛЕННЫЕ СПЛАЙНЫ И ИХ ИСПОЛЬЗОВАНИЕ ДЛЯ СГЛАЖИВАНИЯ ВЫБРОСОВ И ИЗЛОМОВ ИНТЕРПОЛЯНТА

© 2021 В.А. Коднянко

Сибирский федеральный университет

(660074 Красноярск, ул. Академика Куренского, д. 26А)

E-mail: kowlad@rambler.ru

Поступила в редакцию: 18.07.2020

Сформулирован и предложен метод построения направленного кубического сплайна для набора точек на плоскости. Проведено сравнение сплайна с B -сплайном Шёнберга, сплайнами Акимы и Катмулла—Рома. Показано, что для неравноотстоящих точек в сравнении с B -сплайном он дает значительно меньшие выбросы и практически лишен сильных изломов, которые свойственны сплайнам Акимы. Сплайн не дает петель и осцилляций, которые являются характерным недостатком параметрических сплайнов, в частности, эрмитовых, к числу которых относится сплайн Катмулла—Рома. Предложен быстрый метод оптимизации направляющего коэффициента сплайна, цель которой состоит в минимизации разрывов второй производной функции в ее промежуточных точках. Приведен пример оптимизации направленного сплайна третьего порядка. Также предложен направленный сплайн четвертого порядка, который лишен изломов. Сформулирован метод оптимизации направленного сплайна четвертого порядка, изложен алгоритм его оптимизации. Критериями оптимизации являются длина сплайна и наименьшее расстояние между его глобальными максимумом и минимумом. Показано, что в сравнении с сплайном Шёнберга направленный сплайн четвертого порядка имеет меньшие выбросы. Предложен метод автоматического притупления острых пиков кривых, который можно применять ко всем типам сплайнов.

Ключевые слова: сплайн, сплайн Шёнберга, сплайн Акимы, направленный сплайн.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Коднянко В.А. Направленные сплайны и их использование для сглаживания выбросов и изломов интерполянта // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2021. Т. 10, № 1. С. 5–19. DOI: 10.14529/cmse210101.

Введение

При проведении вычислительных или натуральных экспериментов однофакторную связь между входной и выходной величинами обычно определяют посредством вычислений либо измерений. Результатом являются данные в виде набора точек на плоскости

$$(x_0, y_0), (x_1, y_1), \dots, (x_{n-1}, y_{n-1}), (x_n, y_n), x_{i-1} < x_i, i = 1, 2, \dots, n. \quad (1)$$

На его основе зачастую необходимо получить плавную кривую, которая должна проходить через экспериментальные точки.

Задачу можно решить посредством использования интерполяционных методов [1–3]. Для наборов с малым числом точек удовлетворительные результаты дают интерполяционные методы Лагранжа, Ньютона, Стирлинга [2–4]. Однако, с увеличением количества точек набора (1) интерполяционные полиномы в области крайних точек дают осцилляции недопустимо большой амплитуды [4].

Этого недостатка лишены сплайны [5–10, 20–29], которые находят широкое использование в практических приложениях, как-то: при численном решении нелинейных уравнений, сглаживании функций, сжатии и восстановлении графических изображений, ряде других применений. Наиболее известным среди них является B -сплайн Шёнберга [5, 17]. Сплайн обеспечивает безупречную гладкость интерполянта для равноотстоящих точек x_i ($i = 0, 1, \dots, n$), для которых

$$h_i = x_i - x_{i-1} = h = const. \quad (2)$$

Если (2) не выполняется, то плавность кривых, как правило, нарушается. В таких случаях B -сплайн может давать значительные осцилляции кривой (так называемые «выбросы»), которые имеют место в области отрезков с малыми h_i .

Избежать выбросов B -сплайна позволяют сплайн Акимы [11, 16] и эрмитовы сплайны, частным случаем которых является применяемый в графической геометрии параметрический сплайн Катмулла—Рома [12, 13]. Однако и эти сплайны имеют свои недостатки. Так, сплайн Акимы часто дает неприемлемые изломы графика функции в узловых точках, которые видны на рис. 1 (кривая 2).

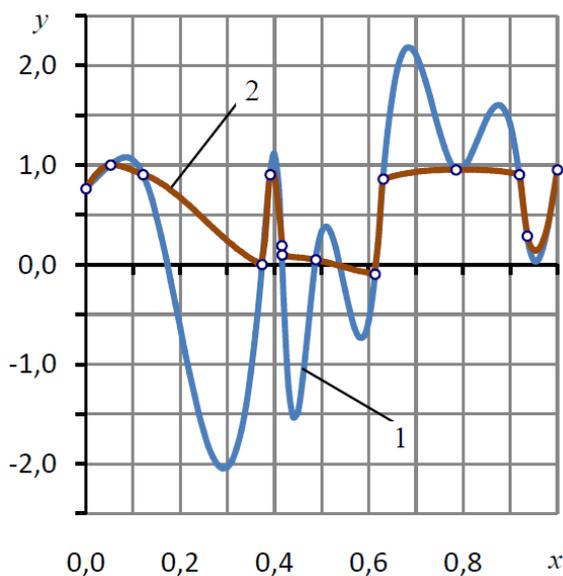


Рис 1. Графики B -сплайна (1) и сплайна Акимы (2)

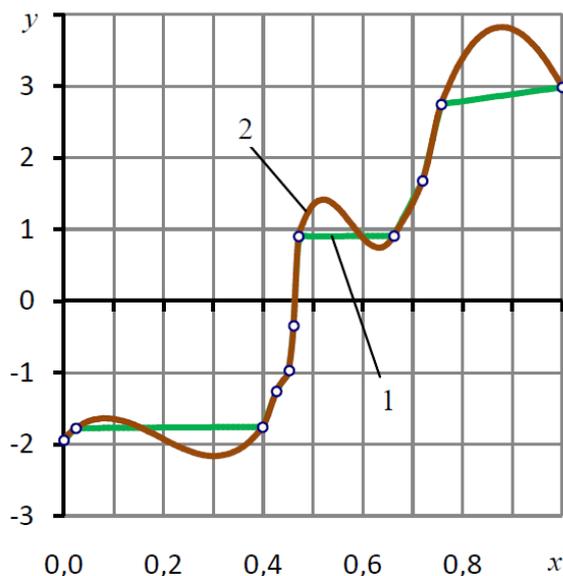


Рис 2. Кривые B -сплайна (1) и сплайна Катмулла—Рома (2)

Более гладкие переходы между соседними полиномами дают сплайны Катмулла—Рома (рис. 2). Однако при малых h_i они могут образовывать петли в узловых точках. Кроме того, на параметрической зависимости $x = x(t)$ которая в силу (1), очевидно, должна быть монотонно-возрастающей, могут появляться локальные экстремумы, что свидетельствует о явных недостатках алгоритма интерполяции при обработке наборов данных, содержащих точки, находящиеся на относительно малом расстоянии друг от друга.

Таким образом, актуальной является задача разработки сплайна, который характеризуется меньшими выбросами B -сплайнов и существенно менее выраженными изломами сплайнов Акимы. Наряду с этим поставлена задача создания такой модели сплайна, которая не уступала бы по скорости вычислений алгоритму сплайна Акимы при изменении одной или нескольких точек, что в сравнении с B -сплайном является характерным преимуществом данного сплайна.

Статья организована следующим образом. В разделе 1 рассмотрена методика построения направленного кубического сплайна (НСЗ-сплайна). В разделе 2 предложен метод и приведен пример оптимизации такого сплайна. Раздел 3 посвящен построению направленного сплайна четвертого порядка (НС4-сплайна). В разделе 4 изложена методика оптимизации направленного сплайна четвертого порядка по критерию минимума длины кривой и минимума расстояния между ее глобальными максимумом и минимумом. В раз-

деле 5 предложен метод притупления острых пиков сплайнов. Методика может быть применена ко всем типам рассмотренных сплайнов. В заключении приводится краткая сводка результатов, полученных в рамках данного исследования, и указаны направления дальнейших исследований.

Ниже рассмотрена методика построения и алгоритм оптимизации такого сплайна.

1. Построение направленного кубического сплайна (НСЗ-сплайна)

На каждом отрезке $[x_{i-1}, x_i]$ будем представлять сплайн $S(x)$ полиномом третьей степени

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, i = 1, 2, \dots, n. \quad (3)$$

На стыках соседних отрезков $[x_{i-1}, x_i]$ для полиномов (3) потребуем выполнения условий непрерывности сплайна и ее первой производной в точках (1)

$$S_i(x_{i-1}) = S_{i-1}(x_{i-1}), \quad (4)$$

$$S'_i(x_{i-1}) = S'_{i-1}(x_{i-1}). \quad (5)$$

Используя (3)–(5), получим

$$\begin{cases} a_i = y_i, \\ a_i - h_i b_i + h_i^2 c_i - h_i^3 d_i = y_{i-1}, \\ b_i - 2h_i c_i + 3h_i^2 d_i = b_{i-1}. \end{cases} \quad (6)$$

Если считать b_i известными, то (6) позволяет получить систему уравнений относительно неизвестных коэффициентов c_i, d_i

$$\begin{cases} c_i - h_i d_i = \frac{b_i - u_i}{h_i}, \\ 2c_i - 3h_i d_i = \frac{b_i - b_{i-1}}{h_i}, \end{cases} \quad (7)$$

где

$$u_i = \frac{y_i - y_{i-1}}{h_i}.$$

Решив (6), найдем

$$c_i = \frac{2b_i + b_{i-1} - 3u_i}{h_i}, d_i = \frac{b_i + b_{i-1} - 2u_i}{h_i^2}. \quad (8)$$

В простейшем случае для крайних отрезков $[x_0, x_1]$ и $[x_{n-1}, x_n]$ можно воспользоваться производной линейного интерполяционного сплайна, положив

$$b_0 = u_1, b_n = u_n.$$

Эти же коэффициенты можно определить при помощи построенного по трем точкам $(x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1})$ интерполяционного полинома Лагранжа [2]

$$f_i(x) = y_i + \frac{[u_i h_{i+1} + u_{i+1} h_i + (u_{i+1} - u_i)(x - x_i)](x - x_i)}{h_{i+1} + h_i}, i = 1, 2, \dots, n - 1,$$

$$b_0 = f'_1(x_0) = u_1 - \frac{h_1(u_2 - u_1)}{h_1 + h_2}, b_n = f'_{n-1}(x_n) = u_n + \frac{h_n(u_n - u_{n-1})}{h_{n-1} + h_n}.$$

Для промежуточных отрезков $[x_{i-1}, x_i]$ будем использовать формулу

$$b_i = \alpha u_i + (1 - \alpha) u_{i+1}, \quad (9)$$

где коэффициент $\alpha \in [0, 1]$.

Из (4)–(9) следует, что на отрезке $[x_{i-1}, x_i]$ функция $S_i(x)$ полностью определяется только тремя точками, в то время как локальность сплайна Акимы определяется пятью точками [11], а B -сплайн всеми точками набора (1). Это свойство характеризует лучшее сравнительное быстродействие рассматриваемого метода при коррекции коэффициентов (3) на случай изменения отдельных точек набора (1).

Разделенная разность u_i представляет собой угловой коэффициент соответствующего сегмента линейного интерполяционного сплайна, который показан на рис. 3.

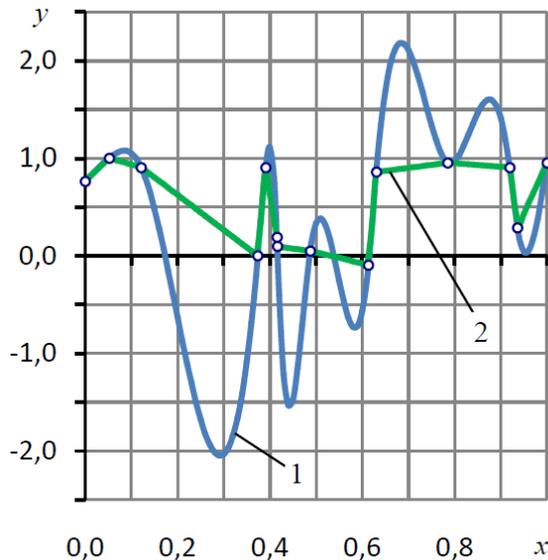


Рис 3. B -сплайн (1) и линейный интерполяционный сплайн (2)

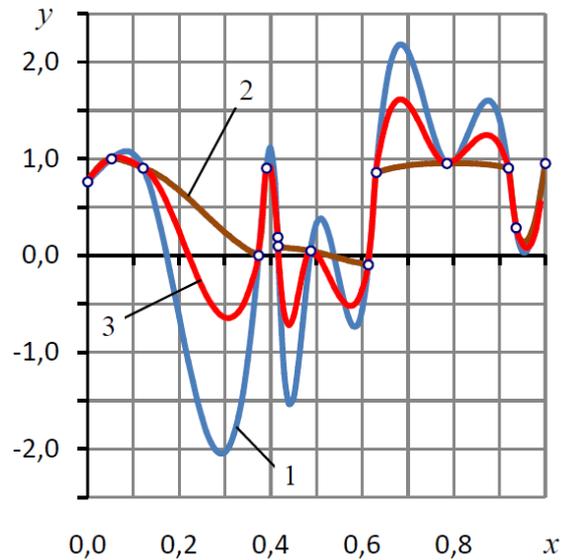


Рис 4. B -сплайн (1), сплайн Акимы (2) и направленный сплайн (3)

Очевидно, (9) является угловым коэффициентом касательной в точке (x_i, y_i) сопряжения сегментов сплайна, находящихся по обе стороны от нее. Варьируя коэффициент α , можно отрегулировать положение касательных в промежуточных точках сплайна, направляя его так, чтобы выбросы были бы минимальными, а изломы не слишком заметными. Назовем такой сплайн *направленным*, а коэффициент α *направляющим*.

В простейшем случае можно положить $\alpha = 0,5$, то есть считать, что направляющая касательная сплайна в промежуточных узловых точках должна занимать среднее положение относительно соседних отрезков линейного интерполяционного сплайна.

На рис. 4 показан пример интерполяции набора данных при помощи B -сплайна (1), сплайна Акимы (2) и направленного сплайна (3).

Видно, что, в отличие от сплайна Акимы, направленный сплайн не имеет видимых изломов, чего для многих практических применений оказывается достаточным основанием для принятия решения об удовлетворительной аппроксимации исходной табличной функции подобными сплайнами. В сравнении с B -сплайном менее выражены и выбросы сплайна.

2. Оптимизация НСЗ-сплайна

Оптимизация преследует цель сделать менее заметными изломы направленного сплайна, которые определяются абсолютной величиной разности значений вторых производных в точке сопряжения соседних полиномов. Величину такого разрыва в промежуточных точках можно определить формулой

$$D_i(\alpha) = \frac{\varepsilon}{2} |S_i''(x_{i-1}) - S_{i-1}''(x_{i-1})| = |c_i - c_{i-1} - 3d_i|, (i = 1, 2, \dots, n - 1). \quad (10)$$

Вычислительный эксперимент показал, что в большинстве случаев при крайних значениях направляющего коэффициента $\alpha = 0$ и $\alpha = 1$, когда углы касательной к кривой в промежуточных точках и одного из отрезков линейного сплайна совпадают, направленный сплайн обычно имеет не только выраженные изломы, но и большие выбросы. При промежуточных значениях α эти недостатки менее заметны. Следовательно, существуют такие значения α , при котором эти недостатки будут наименее выражены.

Суть оптимизации заключается в отыскании такого $\alpha = \alpha_{opt}$, при котором наибольший разрыв

$$D(\alpha) = \max D_i(\alpha), i = 1, 2, \dots, n - 1 \quad (11)$$

кажется минимальным.

Установлено, что в большинстве случаев $D(\alpha)$ является кусочно-линейной функцией с единственной точкой излома. Однако нередко случаи функций с несколькими изломами.

Найти значение α_{opt} можно одним из методов минимизации унимодальных функций [14, 15, 17]. Однако можно воспользоваться свойством кусочной линейности $D(\alpha)$ и на этом основании предложить более быстрый алгоритм.

Методика решения задачи и описание алгоритма нахождения α_{opt} приведены ниже. Используются величины типа $T = (T.x, T.y, T.z)$, где $T.x$, $T.y$ — абсцисса и ордината точки, $T.z$ — значение производной функции $D(\alpha)$ в этой точке.

Шаг 1. Зададим достаточно малое число ε — точность определения α_{opt} , а также границы интервала $A.x = 0$; $B.x = 1$.

Шаг 2. Найдем значение функции $A.y = D(A.x)$ и $A.z = [D(A.x + \varepsilon) - A.y]/\varepsilon$ на левом конце интервала поиска. Определим $B.y = D(B.x)$ и $B.z = [B.y - D(B.x - \varepsilon)]/\varepsilon$ — аналогичные им величины на правом конце интервала поиска.

Шаг 3. По этим точкам и производным построим прямые, первая из которых проходит через точку $(A.x, A.y)$, вторая через $(B.x, B.y)$. Нетрудно показать, что абсциссу точки пересечения этих прямых можно найти по формуле

$$x = \frac{B.y - A.y - B.zB.x + A.zA.x}{A.z - B.z} \quad (12)$$

Вычислим абсциссу $C.x = x + \varepsilon/3$, сдвинутую вправо от x на величину меньшую ε . Это необходимо для того чтоб точка x попала в диапазон $[A.x, C.x]$.

Шаг 4. Если $|C.x - A.x| < \varepsilon$ или $|C.x - B.x| < \varepsilon$, то точка минимума $\alpha_{opt} = x$ найдена и алгоритм заканчивает работу, иначе находим аналогичные значения $C.y$ и $C.z$. Если $C.z$ и $A.z$ числа одного знака, то полагаем $B = C$, иначе $A = C$ и переходим к шагу 2 для выполнения новой итерации.

Продemonстрируем работу алгоритма на примере оптимизации функции с двумя изломами.

Для этого на шаге 1 зададим точность поиска $\varepsilon = 10^{-3}$. На шаге 2 получим $A.x = 0$; $A.y = 1,2$; $A.z = -1,7$; $B.x = 0,999$; $B.y = 0,6$; $B.z = 1,7$. Разные знаки производных $A.z$ и $B.z$ указывают на то, что функция $D(\alpha)$ строго унимодальна [18] и, следовательно, ее минимум находится внутри отрезка.

На шаге 3 по формуле (12) найдем $x = 0,644$; $C.x = 0,645$. Поскольку на шаге 4 ни одно из его условий не выполнилось, то вычислим $C.y$ и $C.z = 0,8 > 0$. Это означает, что справа от точки x функция возрастает, следовательно, минимум находится слева от x . Положим $B = C$ и поиск минимума функции продолжим на отрезке $[0; 0,645]$ меньшей длины.

На новой итерации получим $x = 0,425$; $C.x = 0,426$; $C.z = 0,8 > 0$ и новый отрезок $[0; 426]$.

На следующей итерации найдем $x = 0,425$; $C.x = 0,426$. Теперь условие $|C.x - B.x| < \varepsilon$ выполнилось. Это значит, что минимум функции находится в точке $\alpha_{opt} = x = 0,425$.

Очевидно, число итераций не превышает $k + 1$, где k — число изломов функции $D(\alpha)$. В частности, данная задача решена за три итерации при двух изломах минимизируемой функции.

Приведем таблицу значений функции $y(x) = \sin x$, использованной в работе [19] для оценки погрешности B -сплайна и сплайна Акимы. Сплайны построены на множестве из 14 случайных точек. Вычисления проведены для равноотстоящих узлов. Для сравнения в таблицу добавлены данные по направленному сплайну. В таблице в колонках Δ_1 , Δ_2 , Δ_3 приведены разности между точным значением функции $y(x)$ и значениями, которые получены с помощью B -сплайна (Δ_1), сплайна Акимы (Δ_2) и направленного сплайна (Δ_3), соответственно.

Таблица
Сравнительные данные по вычислительной погрешности сплайнов

x	$y(x)$	Δ_1	Δ_2	Δ_3
0,00	0,0000000	0,000E+00	0,000E+00	0,000E+00
0,05	0,0499792	7,819E-08	-6,100E-05	3,529E-05
0,10	0,0998334	3,189E-08	-1,010E-05	-5,012E-06
0,15	0,1494381	1,433E-08	7,158E-06	9,359E-06
0,20	0,1986693	8,864E-09	2,348E-05	-5,406E-06
0,25	0,2474040	2,986E-08	-1,177E-05	3,179E-06
0,30	0,2955202	8,974E-11	-1,624E-06	6,155E-07
0,35	0,3428978	2,180E-08	6,243E-06	-3,199E-06
0,40	0,3894183	2,905E-08	-8,113E-06	4,885E-06
0,45	0,4349655	-2,711E-20	-2,711E-20	-5,421E-20
0,50	0,4794255	4,830E-10	5,141E-06	-4,438E-06
0,55	0,5226872	8,824E-08	-2,120E-06	2,857E-06
0,60	0,5646425	-6,664E-08	-6,810E-06	2,393E-06
0,65	0,6051864	-1,044E-07	4,300E-06	-1,165E-06
0,70	0,6442177	8,651E-07	3,549E-06	-2,821E-06
0,75	0,6816388	-2,716E-06	-1,707E-06	1,296E-06

Более высокую точность показал B -сплайн. Среди двух последних лучшие результаты дал направленный сплайн. Это неочевидный результат, поскольку ожидалось, что якобы имеющий преимущества при интерполяции монотонных функций, а именно таковой в данном примере является функция $y(x)$, сплайн Акимы должен был бы демонстрировать лучшие показатели не только в сравнении с направленным сплайном, но и с B -сплайном. Однако в данном случае эти ожидания не оправдались.

Изложенная идея позволяет расширить рамки подхода к построению направленного нового сплайна, который будет лишен изломов, и допускает возможность его оптимизации с целью подавления выбросов. Примером может служить направленный сплайн четвертой степени, методика построения которого изложена ниже.

3. Построение направленного сплайна четвертой степени (НС4-сплайна)

На каждом отрезке $[x_{i-1}, x_i]$ будем представлять сплайн $S(x)$ полиномом четвертой степени

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_{i-1})^3 + e_i(x - x_i)^4, i = 1, 2, \dots, n. \quad (13)$$

На стыках соседних отрезков $[x_{i-1}, x_i]$ для полиномов (3) потребуем выполнения условий непрерывности сплайна и ее первой и второй производных в точках (1)

$$S_i(x_{i-1}) = S_{i-1}(x_{i-1}), \quad (14)$$

$$S'_i(x_{i-1}) = S'_{i-1}(x_{i-1}), \quad (15)$$

$$S''_i(x_{i-1}) = S''_{i-1}(x_{i-1}). \quad (16)$$

Полагая по-прежнему b_i известными и, используя (14), получим

$$c_i = d_{i-1}h_i - e_i h_i^2 + w_i, \quad (17)$$

$$e_i = \frac{w_i - c_i + d_i h_i}{h_i^2}, \quad (18)$$

где

$$w_i = \frac{b_i - u_i}{h_i}.$$

Используя (15), найдем

$$2c_i + 4e_i h_i^2 = v_i - \frac{3h_{i-1}^2}{h_i} d_{i-1}, \quad (19)$$

где

$$v_i = \frac{b_i - b_{i-1}}{h_i}.$$

Подставив (18) в (19), запишем

$$c_i = 2d_i h_i + \frac{3h_{i-1}^2}{2h_i} d_{i-1} + 2w_i - \frac{v_i}{2}, \quad (20)$$

$$e_i h_i^2 = \frac{v_i}{2} - d_i h_i - \frac{3h_{i-1}^2}{2h_i} d_{i-1} - w_i. \quad (21)$$

Условие (16) дает зависимость

$$c_i + 6e_i h_i^2 = c_{i-1} + 3d_{i-1} h_i, \quad (22)$$

Подставив (20), (21) в (22), найдем

$$8d_i h_i + d_{i-1} \left(15 \frac{h_{i-1}^2}{h_i} + 10h_{i-1} \right) + \frac{3h_{i-2}^2}{h_{i-1}} d_{i-2} = 5v_i + v_{i-1} - 4(2w_i + w_{i-1}). \quad (23)$$

Сдвинув в (23) индекс, получим рекуррентные формулы

$$\mu_i d_{i+1} + \eta_i d_i + \lambda_i d_{i-1} = \omega_i, i = 1, 2, \dots, n - 1, \quad (24)$$

где

$$\mu_i = 8h_{i+1}, \eta_i = 5h_i \left(3 \frac{h_i^2}{h_{i+1}} + 2 \right), \lambda_i = \frac{3h_{i-1}^2}{h_i}, \omega_i = 5v_{i+1} + v_i - 4(2w_{i+1} + w_i).$$

Формулы (24) представляют трехдиагональную систему линейных алгебраических уравнений относительно неизвестных коэффициентов d_i , которая с учетом очевидных краевых условий $d_0 = 0, d_n = 0$ может быть решена методом прогонки [12]. Далее коэффициенты c_i, e_i можно найти по формулам (17), (18).

4. Оптимизация НС4-сплайна

При оптимизации НС3-сплайна использована однопараметрическая процедура. Это объясняется тем, что многопараметрическая оптимизация в пределе дает B -сплайн Шёнберга, что влечет утрату достоинств НС3. Сплайн НС4 лишен изломов, поэтому его оп-

тимизация сводится к только максимальному подавлению выбросов. В этом процессе могут быть задействованы все направляющие коэффициенты сплайна α_i , при помощи которых вычисляются коэффициенты HC4-сплайна

$$b_i = \alpha_i u_i + (1 - \alpha_i) u_{i+1}, i = 1, 2, \dots, n - 1. \quad (25)$$

В качестве критериев оптимальности данного сплайна использовали

- длину L сплайна,
- разность R между его глобальным максимумом и глобальным минимумом.

Критерий L определяется суммой длины сегментов сплайна и может быть вычислен при помощи известной формулы [23]

$$L = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \sqrt{1 + [s'_i(x)]^2} dx,$$

где

$$S'_i(x) = b_i + 2c_i(x - x_i) + 3d_i(x - x_{i-1})^2 + 4e_i(x - x_i)^3. \quad (26)$$

Для вычисления критерия R использована формула (26), а также формула

$$S''_i(x) = 2c_i + 6d_i(x - x_i) + 12e_i(x - x_i)^2. \quad (27)$$

Используя (26), находили корни уравнения

$$S'_i(x) = 0,$$

их тип контролировали с помощью (27).

Таким образом, как критерий L , так и критерий R являются функциями многих переменных

$$K = K(\alpha), \quad (28)$$

где

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{n-1}).$$

Очевидно, B -сплайн Шёнберга является частным случаем HC4-сплайна. Следовательно, с позиций минимума указанных критериев оптимальный сплайн не может быть хуже сплайна Шёнберга.

В процессе минимизации критериев придерживались требования поддержания такой формы сплайнов, изогомертия которых соответствовала бы сплайну Шёнберга [25].

Расчеты HC4-сплайна показали, что без принятия мер сплайн может терять указанную форму.

Среди причин потери формы выделены следующие:

- сплайн может иметь «выпячивания» отдельных фрагментов кривой,
- на кривой могут появляться новые локальные экстремумы,
- могут также появляться новые точки смены кривизны сплайна.

Данные недостатки связаны с появлением новых локальных экстремумов, как самого сплайна, так и его первой и второй производных, а также новых смен знака его третьей производной. При оптимизации варианты таких сплайнов отбраковывались.

Численные эксперименты позволили установить тот факт, что функция (28) является многоэкстремальной, то есть имеет множество локальных минимумов, среди которых следует найти глобальный минимум, который соответствует сплайну оптимальной формы.

Так, при $n = 12$, для которого проводилось большинство экспериментов, пришлось бы отыскивать глобальный минимум функции 11 переменных, что представляется задачей чрезвычайной сложности. В общем случае трудности получения решения такой задачи ставят под сомнение ценность практического использования обсуждаемого сплайна. Выход из положения был найден в использовании упомянутых свойств сплайна, которые продиктованы жесткими условиями сохранения его формы.

Методика минимизации критериев состоит в следующем. Изначально устанавливается стартовое состояние, для которого выбирается вектор α , все компоненты которого полагаются равными 0,5, и вычисляется стартовый НС-сплайн. Далее назначается шаг τ малой длины и проводится переборная однопараметрическая минимизация по каждой компоненте вектора α , для которой процесс начинается со стартового состояния. Полученные таким образом лучшие покоординатные сплайны дают множество α -векторов, число которых $m < n$. Наблюдения показали, что через фильтр жестких требований сохранения формы сплайна обычно проходит не более половины начальных стартов. Так, например, для $n = 12$ обычно $m < 7$. Следующим шагом является аналогичная однопараметрическая оптимизация, где в качестве стартов последовательно используются векторы, прошедшие фильтр первого шага. При этом количество новых векторов, которые прошли фильтр сохранения формы также невелико и оно обычно меньше аналогичного количества предыдущего шага. Рекурсия ведется до тех пор, пока отфильтрованные векторы не дадут новые векторы для продолжения процесса. Результатом оптимизации будет сплайн с наименьшим значением критерия K . Расчеты показали, что, например, для $n = 12$ обычно требуется сформировать сплайн и вычислить его характеристики 1000–2000 раз.

На рис. 5 показаны кривые сплайнов, которые дают типичную картину формы оптимального по критерию длины НС4-сплайна. Сплайн обычно занимает среднее положение между сплайном Шёнберга и НС3 ближе к первому.

Приведем характерный пример процесса оптимизации НС4. Длина сплайна Шёнберга 19,70. Исходный НС4 имеет несколько большую длину — 19,90. Оптимизированный сплайн имеет длину 16,82, что на 14,6% меньше длины сплайна Шёнберга. При проведении оптимизации алгоритм улучшал результат 181 раз. Для решения задачи потребовалось провести вычисление сплайна 1312 раз. Наименьшую длину имел НС3. При длине 11,58 он почти в два раза короче сплайна Шёнберга. Тысячи вычислительных экспериментов, поставленных для $n = 12$, показали, что оптимизированный НС4 короче сплайна Шёнберга на 5–50%, а НС3 — в 1,5–3 раза.

Примеры процессов оптимизации можно наблюдать в динамике на видео по ссылкам [30–32].

Исследования показывают, НС4 позволяет ослабить проявление выбросов сплайна Шёнберга. Однако лучшим в этом отношении следует признать СПЗ, на котором выбросов нет, а изломы проявляются весьма слабо.

В процессе изучения свойств предложенных сплайнов обнаружено, что кроме выбросов и изломов сплайны могут иметь острые пики локальных максимумов и минимумов, которые в ряде случаев также следует рассматривать как недостатки интерполяции. Ниже предложен и описан способ, который позволять притупить пики экстремумов.

5. Методика притупления пиков сплайнов

Пусть необходимо притупить локальный минимум какого-либо из рассмотренных сплайнов и пусть (x_c, y_c) точка такого минимума. Рассмотрим также точки перегиба (x_a, y_a) и (x_b, y_b) , расположенные по обе стороны от экстремума.

Для сглаживаемого экстремума сплайна введем добавочную функцию

$$g(x) = G(x - x_a)^k(x_b - x)^m, G \geq 0, k > 2, m \geq 2. \quad (29)$$

Для нее имеют место очевидные краевые условия

$$g'(x_a) = g'(x_b) = 0, g''(x_a) = g''(x_b) = 0.$$

Следовательно, на краях интервала функция не привносит изломов и способствует притуплению сплайна в области его минимума.

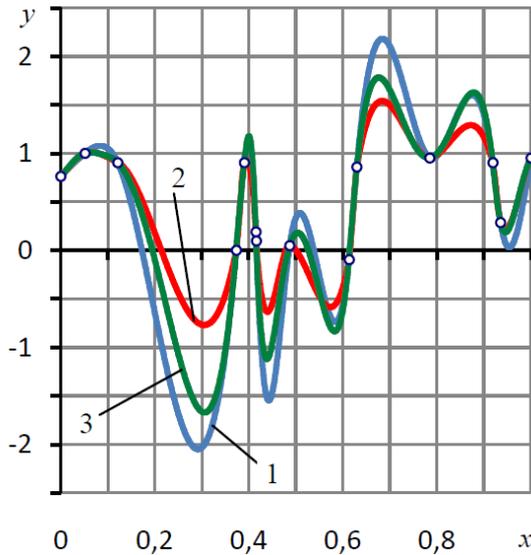


Рис 5. B-сплайн (1), HC3-сплайн (2), HC4-сплайн (3)

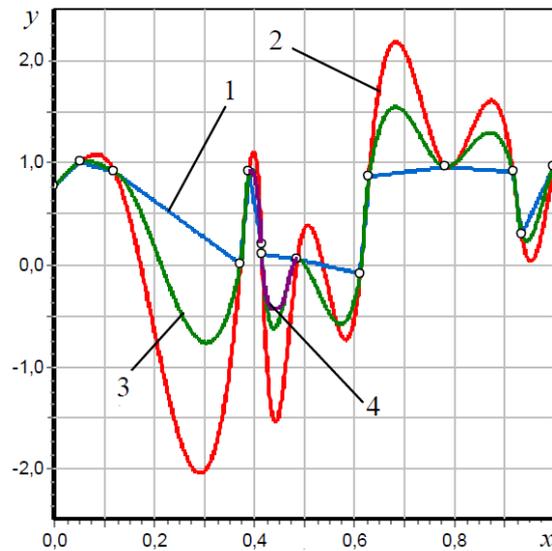


Рис 6. Линейный интерполяционный сплайн (1), B-сплайн (2), HC3 (3) и HC3 с притупленными экстремумами (4)

Максимум этой функции имеет место при

$$g'(x_c) = G(x_c - x_a)^{k-1}(x_b - x_c)^{m-1}[k(x_b - x_c) - m(x_c - x_a)] = 0, \quad (30)$$

откуда следует

$$m = tk, \quad t = \frac{(x_b - x_c)}{(x_c - x_a)}.$$

Если $t < 1$, то примем $m = 2 + \varepsilon$, где ε — малое число. Тогда $k = m/t > 2$. Иначе при $t \geq 1$ примем $k = 2 + \varepsilon$. Тогда $m = kt > 2$.

Обратимся к линейному интерполяционному сплайну

$$L(x) = L_a x + L_b, L(x_a) = y_a, L(x_b) = y_b,$$

где

$$L_a = \frac{y_b - y_a}{x_b - x_a}, L_b = \frac{x_b y_a - x_a y_b}{x_b - x_a}.$$

Для точки рассматриваемого минимума сплайна имеем

$$G(x_c - x_a)^k(x_b - x_c)^m = \sigma[L(x_c) - y_c], \quad (31)$$

где σ — коэффициент притупления пика.

При $\sigma = 0$ добавка $g(x)$ отсутствует, при $\sigma = 1$ суммарный сплайн $q(x) = s(x) + g(x)$ коснется отрезка прямой $L(x)$, что является предельным случаем, поскольку при $\sigma > 1$ суммарный сплайн получит новый экстремум-максимум, по обе стороны которого будет образовано два новых минимума. Такие режимы не обеспечивают сохранения формы сплайна, поэтому $0 \leq \sigma \leq 1$.

Для сохранения формы суммарный сплайн кроме того должен иметь кривизну одного знака до $\sigma = \sigma_{max} < 1$. Данная величина может быть определена из условия

$$q''(x_c) = 0.$$

Выполнив дифференцирование функции $q(x)$, найдем

$$s''(x_c) + G(k + m)G(x_c - x_a)^{k-1}(x_b - x_c)^{m-1} = 0.$$

Отсюда

$$G = \frac{s''(x_c)}{(k + m)(x_c - x_a)^{k-1}(x_b - x_c)^{m-1}}. \quad (32)$$

Из (31), (32) вытекает, что

$$\frac{\sigma_{max}[L(x_c) - y_c]}{(x_c - x_a)^{k-1}(x_b - x_c)^{m-1}} = \frac{s''(x_c)}{(k+m)(x_c - x_a)^{k-1}(x_b - x_c)^{m-1}}. \quad (33)$$

Отношение (33) позволяет определить предельно допустимую величину σ

$$\sigma_{max} = \frac{s''(x_c)(x_c - x_a)(x_b - x_c)}{(k+m)[L(x_c) - y_c]}.$$

Аналогично проводится притупление локальных максимумов.

Решение о подлежащих притуплению экстремумах, может быть принято, как в автоматизированном, так и автоматическом режимах. При автоматическом принятии решения в качестве критерия оценки экстремума можно использовать величину $k_e = s''(x_c)$.

Приведем описание методики оценивания экстремумов, подлежащих автоматическому притуплению. В качестве ориентира будем использовать НСЗ, график которого дан на рис. 4. При визуальной оценке сплайна можно выделить два экстремума, которые необходимо притупить. Один из них является максимумом с абсциссой $x = 0,395$, второй соседним минимумом с абсциссой $x = 0,440$. Для первого $k_e = -2920$, для второго $k_e = 1871$. Следующий за ними по убыванию абсолютной величины критерия k_e является крайний правый минимум с абсциссой $x = 0,946$, для которого $k_e = 1123$. Оценим его как экстремум, не требующий притупления. Данный простой анализ позволяет в первом приближении сформулировать следующую методику автоматической экспертной оценки: экстремумы, которые должны быть притуплены автоматически, должны удовлетворять условию $|k_e| > 1200$, если это не противоречит приведенным выше условиям сохранения формы сплайна.

На рис. 6 в качестве примера притупления экстремумов показаны линейный интерполяционный сплайн, B -сплайн, НСЗ-сплайн и два фрагмента НСЗ-сплайна, на котором экстремумы притуплены автоматически в соответствии с описанной выше методикой. При вычислениях принята величина коэффициента притупления $\sigma = 0,75\sigma_{max}$. Методика притупления острых пиков кривых может быть применена к сплайнам любого типа.

Заключение

В статье сформулирован и предложен простой метод построения кубического сплайна для набора точек на плоскости. Проведено сравнение сплайна с B -сплайном Шёнберга и сплайнами Акимы и Катмулла—Рома. Показано, что для неравноотстоящих точек, при которых обычно проявляются недостатки названных сплайнов, в сравнении с B -сплайном он дает значительно меньшие осцилляции. Сплайн с таким набором точек практически лишен сильных изломов, которые свойственны сплайнам Акимы. Он не дает петель и осцилляций, которые являются характерным недостатком параметрических сплайнов, в частности, эрмитовых, к числу которых относится сплайн Катмулла—Рома. Предложен метод оптимизации направляющего коэффициента сплайна, цель которой состоит в минимизации разрывов второй производной функции в ее промежуточных точках. Также предложен направленный сплайн четвертого порядка, который лишен изломов и в сравнении со сплайном Шёнберга имеет меньшие выбросы. Предложен метод притупления острых пиков кривых, который можно применять ко всем типам сплайнов. Приведенные в статье численные и графические результаты получены на основе предложенных алгоритмов, реализованных при помощи разработанного автором программного обеспечения в среде визуального программирования Delphi с использованием приложения MS Office Excel.

В ходе дальнейшей работы планируется улучшить реализацию алгоритма оптимизации направленного сплайна четвертого порядка с целью сокращения времени его работы и повышения точности процедуры минимизации критериев оптимальности сплайна.

Литература

1. Powell M.J.D. Approximation Theory and Methods. Cambridge University Press, 1981. 352 p. DOI: 10.1017/CBO9781139171502.
2. Atkinson K.A. An Introduction to Numerical Analysis (2nd ed.). John Wiley and Sons, 1988. 615 p.
3. Watson G.A. Approximation Theory and Numerical Methods. John Wiley, 1980. 229 p.
4. Schatzman M. Numerical Analysis: A Mathematical Introduction. Clarendon Press, 2002. 496 p.
5. Schoenberg I.J. Contributions to the problem of approximation of equidistant data by analytic functions // Quart. Appl. Math. 1946. Vol. 4. P. 45–99, 112–141.
6. Ahlberg J.H., Nilson E.N., Walsh J.L. The Theory of Splines and Their Application. Academic Press, 1967. 296 p.
7. David F., Rogers J., Adams A. Mathematical Elements for Computer Graphics. McGraw-Hill Science / Engineering / Math, 2 edition, 1989. 611 p.
8. Cohen D. Incremental Methods for Computer Graphics. PhD Thesis. Harvard University, 1969.
9. Warnock J.E. A Hidden Surface Algorithm for Computer-Generated Halftone Pictures. Computer Science Department, University of Utah, TR 1–15. 1969.
10. Watkins G.S. A Real-Time Visible Surface Algorithm. Computer Science Department, University of Utah, UTECH-CSC-70-101, 1970.
11. Akima H. A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures // Journal of the ACM. 1970. Vol. 17, no. 4. P. 589–602. DOI: 10.1145/321607.321609.
12. Catmull E., Rom R. A class of local interpolating splines // Computer Aided Geometric Design. 1974. P. 317–326.
13. Barry P.J., Goldman R.N. Recursive evaluation algorithm for a class of Catmull-Rom splines // Computer Graphics. 1988. Vol. 22, no. 4. P. 199–204. DOI: 10.1145/378456.378511.
14. Kiefer J.K. Sequential minimax search for a maximum // P. Am. Math. Soc. 1953. P. 502–506.
15. Brent R.P. Algorithms for Minimization without Derivatives. Dover. 2002. 195 p.
16. Круковец А.С., Горелкин Г.А. Разработка метода интерполяции значений номограммы // Современные научные исследования и инновации. 2015. № 5(2). URL: <http://web.snauka.ru/issues/2015/05/53846> (дата обращения: 14.06.2020).
17. Dobson A.J. An Introduction to Statistical Modelling. Chapman and Hall, London, 1983.
18. Коднянко В.А. О вычислительной избыточности метода дихотомии и условной минимизации унимодальных функций методом экономной дихотомии // Системы и средства информатики. 2019. Т. 29, № 1. С. 164–173. DOI: 10.14357/08696527190113.
19. Ruckdeschel F.R. Basic scientific subroutines. Vol. 2. BYTE/McGRAW-HILL, 1981.
20. Яненко Н.Н., Квасов Б.И. Итерационный метод построения поликубических сплайн-функций // Докл. АН СССР. 1970. Т. 195, № 5. С. 1055–1057.
21. Constantini P., Morandi R. An algorithm for computing shape-preserving cubic spline interpolation to data // Calcolo. 1984. Vol. 21. P. 295–305.
22. Рябенский В.С. Локальные формулы гладкого восполнения и гладкой интерполяции функций по их значениям в узлах неравномерной прямоугольной сетки // Препринт. Ин-т прикл. математики АН СССР. 1974. № 21. 35 с.

23. Dietze S., Schmidt J.W. Determination of shape preserving spline interpolants with minimal curvature via dual programs // J. Approxim. Theory. 1988. Vol. 52, no. 1. P. 43–57.
24. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. М.: Наука, 1980. 352 с.
25. Мирошниченко В.Л. Изогометрические свойства и погрешность аппроксимации взвешенных кубических сплайнов // Вычислительные системы. Новосибирск: ИМ СО РАН, Вып. 154: Сплаины и их приложения. 1995. С. 127–154.
26. Корнейчук Н.П., Бабенко В.Ф., Лигун А.А. Экстремальные свойства полиномов и сплайнов. Киев: Наукова думка. 1992. 304 с.
27. Dzyubenko G.A., Gilewicz J., Shevchuk I.A. New phenomena in coconvex approximation // Analysis Mathematica. 2006. Vol. 32, no. 2. P. 113–121. DOI: 10.1007/s10476-006-0005-x.
28. Стечкин С.Б., Субботин Ю.Н. Сплаины в вычислительной математике. М.: Наука, 1976. 248 с.
29. Волков Ю.С. Новый способ построения интерполяционных кубических сплайнов // Журн. вычисл. матем. и матем. физ. 2004. Т. 44, № 2. С. 231–241.
30. Коднянко В.А. Video DS4-spline optimization 1. URL: <http://smiuk.sfu-kras.ru/kodnyanko/site/science/Video1.mp4> (дата обращения: 12.07.2020).
31. Коднянко В.А. Video DS4-spline optimization 2. URL: <http://smiuk.sfu-kras.ru/kodnyanko/site/science/Video2.mp4> (дата обращения: 12.07.2020).
32. Коднянко В.А. Video DS4-spline optimization 3. URL: <http://smiuk.sfu-kras.ru/kodnyanko/site/science/Video3.mp4> (дата обращения: 12.07.2020).

Коднянко Владимир Александрович, д.т.н, профессор, кафедра стандартизации, метрологии и управления качеством, Сибирский федеральный университет (Красноярск, Российская Федерация)

DOI: 10.14529/cmse210101

DIRECTIONAL SPLINES AND THEIR USE FOR SMOOTHING EJECTIONS AND FRACTURES OF INTERPOLANT

© 2021 V.A. Kodnyanko

Siberian Federal University (Kirensky 26A, Krasnoyarsk, 660074 Russia)

E-mail: kowlad@rambler.ru

Received: 18.07.2020

A method for constructing a directional cubic spline for a set of points on a plane is formulated and proposed. The spline is compared with the Schoenberg B-spline, Akima and Catmull–Rom splines. It is shown that for unequally spaced points, in comparison with the B-spline, it gives significantly lower overshoots and is practically free of strong kinks, which are characteristic of Akima splines. The spline does not give loops and oscillations, which are a characteristic drawback of parametric splines, in particular, Hermitian ones, which include the Catmull–Rom spline. A fast method for optimizing the spline guiding coefficient is proposed, the purpose of which is to minimize the discontinuities of the second derivative of the function at its intermediate points. An example of optimization of a directional third-order spline is given. A fourth-order directional spline, which is free of kinks, is also proposed. The method of optimization of the directional spline of the fourth order is formulated, the algorithm of its optimization is stated. The optimization criteria are the spline length and the smallest distance between its global maximum and minimum. It is shown that, in comparison with the Schoenberg spline, the fourth-order directional spline has smaller outliers. A method for automatic blunting of sharp peaks of curves is proposed, which can be applied to all types of splines.

Keywords: spline, Schoenberg spline, Akima spline, directional spline.

FOR CITATION

Kodnyanko V.A. Directional Splines and Their Use for Smoothing Ejections and Fractures of Interpolant. *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*. 2021. Vol. 10, no. 1. P. 5–19. (in Russian) DOI: 10.14529/cmse210101.

This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 3.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.

References

1. Powell M.J.D. Approximation Theory and Methods. Cambridge University Press, 1981. 352 p. DOI: 10.1017/CBO9781139171502.
2. Atkinson K.A. An Introduction to Numerical Analysis (2nd ed.). John Wiley and Sons, 1988. 615 p.
3. Watson G.A. Approximation Theory and Numerical Methods. John Wiley, 1980. 229 p.
4. Schatzman M. Numerical Analysis: A Mathematical Introduction. Clarendon Press, 2002. 496 p.
5. Schoenberg I.J. Contributions to the problem of approximation of equidistant data by analytic functions. *Quart. Appl. Math.* 1946. Vol. 4. P. 45–99, 112–141.
6. Ahlberg J.H., Nilson E.N., Walsh J.L. The Theory of Splines and Their Application. Academic Press, 1967. 296 p.
7. David F., Rogers J., Adams A. Mathematical Elements for Computer Graphics. McGraw-Hill Science / Engineering / Math, 2 edition, 1989. 611 p.
8. Cohen D. Incremental Methods for Computer Graphics. PhD Thesis. Harvard University, 1969.
9. Warnock J.E. A Hidden Surface Algorithm for Computer-Generated Halftone Pictures. Computer Science Department, University of Utah, TR 1–15. 1969.
10. Watkins G.S. A Real-Time Visible Surface Algorithm. Computer Science Department, University of Utah, UTECH-CSC-70-101, 1970.
11. Akima H. A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures. *Journal of the ACM*. 1970. Vol. 17, no. 4. P. 589–602. DOI: 10.1145/321607.321609.
12. Catmull E., Rom R. A class of local interpolating splines. *Computer Aided Geometric Design*. 1974. P. 317–326.
13. Barry P.J., Goldman R.N. Recursive evaluation algorithm for a class of Catmull-Rom splines. *Computer Graphics*. 1988. Vol. 22, no. 4. P. 199–204. DOI: 10.1145/378456.378511.
14. Kiefer J.K. Sequential minimax search for a maximum. *P. Am. Math. Soc.* 1953. P. 502–506.
15. Brent R.P. Algorithms for Minimization without Derivatives. Dover, 2002. 195 p.
16. Krukovets A.S., Gorelkin G.A. Development of a method for interpolating the values of the nomogram. *Modern scientific research and innovations*. 2015. No. 5(2). URL: <http://web.snauka.ru/issues/2015/05/53846> (accessed: 14.06.2020) (in Russian)
17. Dobson A.J. An Introduction to Statistical Modelling. Chapman and Hall, London, 1983.
18. Kodnyanko V.A. On computational redundancy of the dichotomous search and conditional minimization of unimodal functions by the economical dichotomous search. *Systems and*

- means of informatics. 2019. Vol. 29, no. 1. P. 164–173. DOI: 10.14357/08696527190113 (in Russian)
19. Ruckdeschel F.R. Basic scientific subroutines. Vol. 2. BYTE/McGRAW-HILL, 1981.
 20. Yanenko N.N., Kvasov B.I. An Iterative Method for Constructing Polycubic Spline Functions. Dokl. USSR Academy of Sciences. 1970. Vol. 195, no. 5. P. 1055–1057. (in Russian)
 21. Constantini P., Morandi R. An algorithm for computing shape-preserving cubic spline interpolation to data. Calcolo. 1984. Vol. 21. P. 295–305.
 22. Ryabenky V.S. Local formulas for smooth completion and smooth interpolation of functions by their values at nodes of an uneven rectangular grid. Preprint, Inst. mathematics of the Academy of Sciences of the USSR. IPM. 1974. No. 21. (in Russian)
 23. Dietze S., Schmidt J.W. Determination of shape preserving spline interpolants with minimal curvature via dual programs. J. Approxim. Theory. 1988. Vol. 52, no. 1. P. 43–57.
 24. Zavyalov Yu.S., Kvasov B.I., Miroschnichenko V.L. Methods of spline functions. Moscow, Nauka, 1980. (in Russian)
 25. Miroschnichenko V.L. Isogeometric properties and approximation error for weighted cubic splines. Computing Systems. Novosibirsk: IM SB RAS, 1995. Vol. 154. P. 127–154. (in Russian)
 26. Korneychuk N.P., Babenko V.F., Ligun A.A. Extreme properties of polynomials and splines. Kiev, Naukova Dumka, 1992. 304 p. (in Russian)
 27. Dzyubenko G.A., Gilewicz J., Shevchuk I. A. New phenomena in coconvex approximation. Analysis Mathematica. 2006. Vol. 32, no. 2. P. 113–121. DOI: 10.1007/s10476-006-0005-x.
 28. Stechkin S.B., Subbotin Yu.N. Splines in computational mathematics. Moscow, Nauka, 1976. 248 p. (in Russian)
 29. Volkov Yu.S. A new method for constructing interpolation cubic splines. Computational Mathematics and Mathematical Physics. 2004. Vol. 44, no. 2. P. 231–241. (in Russian)
 30. Kodnyanko V.A. Video DS4-spline optimization 1. URL: <http://smiuk.sfu-kras.ru/kodnyanko/site/science/Video1.mp4> (accessed: 12.07.2020).
 31. Kodnyanko V.A. Video DS4-spline optimization 2. URL: <http://smiuk.sfu-kras.ru/kodnyanko/site/science/Video2.mp4> (accessed: 12.07.2020).
 32. Kodnyanko V.A. Video DS4-spline optimization 3. URL: <http://smiuk.sfu-kras.ru/kodnyanko/site/science/Video3.mp4> (accessed: 12.07.2020).