

P11-2004-187

Н. Д. Дикусар, Ч. Торок*

**КУСОЧНО-КУБИЧЕСКАЯ АППРОКСИМАЦИЯ
В РЕЖИМЕ АВТОМАТИЧЕСКОГО СЛЕЖЕНИЯ**

*Технический университет г. Кошице, Словакия

Дикусар Н. Д., Торок Ч.

P11-2004-187

Кусочно-кубическая аппроксимация в режиме автоматического слежения

В рамках четырехточечных преобразований предложен метод кусочно-кубической аппроксимации с автоматическим обнаружением узлов сегментов по отсчетам кривой в режиме слежения. В качестве локальной аппроксиманты используется модель трехточечного кубического сплайна (TPS). Точечная оценка коэффициента при x^3 (свободного параметра θ) определяется из уравнения модели, а его интервальная оценка уточняется в процессе отслеживания кубического сегмента кривой. Получено аналитическое выражение параметра θ через длину интервала и значения функции и производных в узлах, указывающее на прямую зависимость C^1 -гладкости от точности вычисления оценки θ . Показана устойчивость метода к входным ошибкам. Основными параметрами аппроксимации являются параметры базисных функций, величина дисперсии входных ошибок и шаг квантования. Эффективность метода и алгоритма подтверждены численными расчетами на примерах аппроксимации сложных кривых и реальных данных.

Работа выполнена в Лаборатории информационных технологий ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна, 2004

Dikoussar N. D., Török Cs.

P11-2004-187

Piecewise-Cubic Approximation in Autotracking Mode

A method for piecewise-cubic approximation within the frame of four-point transforms is proposed. The knots of the segments are detected in autotracking mode using a digitized curve. A three-point cubic parametric spline (TPS) is used as a model of a local approximant. A free parameter θ (a coefficient at x^3) is found in a line following mode, using step-by-step averaging. A formula for expression of the free parameter via a length of the segment and values of a function and derivatives in joining points is received. The C^1 -smoothness depends on the accuracy of the θ -estimate. The stability of the method w.r.t. input errors is shown as well. The key parameters of the approximation are: the parameters of the basic functions, the variance of the input errors, and a sampling step. The efficiency of the method is shown by numerical calculations on test examples.

The investigation has been performed at the Laboratory of Information Technologies.

Communication of the Joint Institute for Nuclear Research. Dubna, 2004

ВВЕДЕНИЕ

В численных расчетах для приближения гладких функций часто используют кусочно-многочленную интерполяцию (аппроксимацию). Соответствующие процедуры включены в большинство современных программных пакетов [1]. Обычно для этих целей используют сплайны. Наиболее интенсивно применяются кубические сплайны, обладающие C^1 - и C^2 -гладкостью в узлах – точках стыковки соседних звеньев сплайна. Как правило, применение сплайн-интерполяции приводит к системам линейных уравнений с ленточными матрицами, решение которых находится методом прогонки [2]. Использование таких процедур в системах реального времени связано с трудностями, такими как оптимальный выбор узлов и большая размерность матриц.

Использование эрмитовой кубической интерполяции освобождает от решения системы уравнений [3], но для обеспечения C^1 -гладкости необходимо задавать оценки или значения первых производных в узлах. Оптимальный выбор узлов представляет крайне трудную задачу [4], поэтому на практике узлы часто определяют методом проб и ошибок.

В системах цифровой обработки сигналов [5, 6] потоки данных — отсчеты с цифроаналоговых преобразователей — поступают последовательно за весьма короткие промежутки времени, и оценки основных параметров сигналов (местоположение пика, амплитуда, площадь и т.п.) необходимо определять в масштабе реального времени. Для решения таких задач требуется разработка более эффективных алгоритмов аппроксимации и сглаживания.

В непараметрических методах [7, 8] важным *параметром сглаживания* является протяженность отрезка (*span*), на котором вычисляется точечная оценка. В случае кусочно-кубической аппроксимации этот параметр ассоциируется с длиной промежутка, на котором кривая приближается кубической дугой. Задача автоматического обнаружения границ таких сегментов (узлов) по входным данным остается актуальной. Для ее решения мы применяем метод четырехточечных преобразований (Discrete Projective Transforms — DPT) [9, 10], в котором базисные функции полиномиальной модели могут использовать абсциссы концов интервала аппроксимации в качестве непрерывных параметров.

В работе на основе DPT-методики построена модель локального трехточечного кубического сплайна (TPS) [10] и разработан простой и устойчивый к ошибкам адаптивный метод (алгоритм) обнаружения по входным данным

узлов кубического сегмента и вычисления оценок его параметров в режиме слежения.

Вначале даются основные определения и обозначения, используемые в конструкции модели. Затем выводится формула для вычисления свободного параметра θ . В разд. 3 и 4 рассматриваются метод и алгоритм для обнаружения узлов кубических сегментов в режиме слежения. Анализ устойчивости метода к ошибкам приводится в разд. 5. В разд. 6, 7 дается краткое изложение процедуры сглаживания и примеры результатов работы алгоритма аппроксимации для различных наборов данных.

1. TPS-МОДЕЛЬ

Пусть множество точек

$$\{x_m, f_m\}_{m=1}^N, \quad f_m = f(x_m); \quad x_m < x_{m+1}, \quad 4 \ll N \quad (1)$$

задает временную последовательность отсчетов. Это могут быть точки на кривой с известной зависимостью $f(x) \in C_{[a,b]}$, результаты вычислений по сложным итерационным схемам, отсчеты оцифровки сигнала в приборе или на изображении, временной ряд и т.п. В системах реального времени число отсчетов N часто бывает достаточно большим и неизвестным заранее, а данные образуют временную последовательность (1). Для обработки таких данных обычно применяются рекуррентные методы (рекурсивный МНК, методы авто-регрессии и др.).

Наш подход к решению таких задач использует рекуррентное вычисление оценки параметра и базируется на параметрической кубической модели, содержащей *один свободный* и *три фиксированных* параметра. При известных фиксированных параметрах метод позволяет найти локально-оптимальную оценку аппроксиманты в режиме слежения (без составления и решения системы алгебраических уравнений).

Введем определения и обозначения, использующиеся в конструкции TPS-модели (рис. 1). Три точки (x_α, r_α) , (x_β, r_β) и (x_0, r_0) , $x_\alpha, x_\beta, x_0 \in [a, b]$, $x_i \neq x_j$, $i, j = \alpha, \beta, 0$ будем называть *репером*. Значения r_* берем из данных ($r_* \equiv f_*$) либо вычисляем их оценки ($r_* = \hat{f}_*$). Из ординат репера и текущей точки (x_τ, f_τ) сформируем два вектора $\mathbf{r}_0 = [r_\alpha, r_\beta, r_0]^T$ и $\mathbf{r} = [r_\alpha, r_\beta, f_\tau]^T$. Перенесем начало в точку x_0 : $\alpha = x_\alpha - x_0$, $\beta = x_\beta - x_0$, $\tau = x_\tau - x_0$, $\alpha \neq \beta \neq 0$. Величины α , β и x_0 жестко связаны с вектором \mathbf{r}_0 и используются как параметры базисных функций модели $d_i(\tau; \alpha, \beta)$, $i = 1, 2, 3$ и $Q(\tau; \alpha, \beta)$ (о d_i и Q см. ниже). Для определенности считаем, что $\alpha < \beta$ и $\tau \in [\alpha, \beta] \subseteq [a - x_0, b - x_0]$.

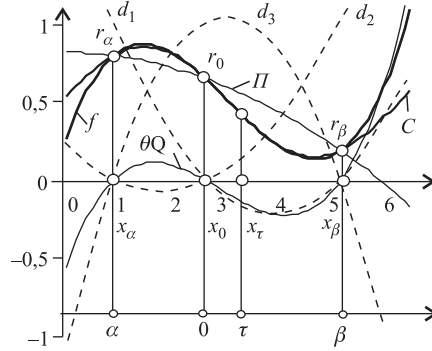


Рис. 1. Конструкция TPS-модели

В этих обозначениях уравнение TPS-модели запишется в виде суммы двух парабол — квадратичной (Π) и кубической (θQ) с параметрами \mathbf{a} , \mathbf{r}_0 и θ :

$$C(\tau; \mathbf{a}, \mathbf{r}_0, \theta) = \Pi(\tau; \mathbf{a}, \mathbf{r}_0) + \theta Q(\tau; \mathbf{a}), \quad \tau \in [a - x_0, b - x_0], \quad (2)$$

где $\mathbf{a} = (\alpha, \beta, x_0)$; θ — свободный параметр, совпадающий с коэффициентом при τ^3 ; Q — «зануляющая» кубическая парабола вида

$$Q = \tau(\tau - \alpha)(\tau - \beta), \quad (3)$$

а $\Pi(\tau; \mathbf{a}, \mathbf{r}_0)$ — «реперная парабола», проходящая через реперные точки \mathbf{r}_0 . С учетом вектора весов $\mathbf{d} = [d_1, d_2, d_3]^T$ ее уравнение принимает следующий вид

$$\Pi(\tau; \mathbf{a}, \mathbf{r}_0) = (\mathbf{r}_0, \mathbf{d}). \quad (4)$$

Функции $d_i(\tau; \alpha, \beta)$, $i = 1, 2, 3$ определяются специальным правилом ангармонического отношения четырех точек $\{\tau\alpha\beta 0\}$ [9,10], порождающего три функции, квадратичные относительно τ и дробно-рациональные относительно α и β :

$$\begin{aligned} d_1 &= \frac{-\tau(\tau - \beta)}{\alpha\gamma}, & d_2 &= \frac{\tau(\tau - \alpha)}{\beta\gamma}, \\ d_3 &= \frac{(\tau - \alpha)(\tau - \beta)}{\alpha\beta}, & \text{где } \gamma &= \beta - \alpha. \end{aligned} \quad (5)$$

Отметим, что параметры α и β входят в конструкцию d_i и Q непрерывным образом, что позволяет использовать их в качестве новых переменных. При известном \mathbf{r}_0 модель (2) содержит один неизвестный параметр — коэффициент θ ($\theta = C'''/6$). В целом TPS-модель зависит от семи параметров \mathbf{a} ,

\mathbf{r}_0 и θ , тогда как (с учетом узлов) стандартная модель $f = \sum_{i=0}^3 c_i x^i$ использует только шесть. Коэффициенты c_i выражаются через параметры TPS-модели следующими формулами

$$\begin{aligned} c_3 &= \theta, & c_2 &= (\alpha\Delta_{\beta 0} - \beta\Delta_{\alpha 0})/(\alpha\beta\gamma) + \theta\alpha\beta, \\ c_1 &= (\beta^2\Delta_{\alpha 0} - \alpha^2\Delta_{\beta 0})/(\alpha\beta\gamma) - \theta(\alpha + \beta), & c_0 &= f_0, \end{aligned} \quad (6)$$

где $\Delta_{\alpha 0} = f_\alpha - f_0$, $\Delta_{\beta 0} = f_\beta - f_0$.

Из (2) и рис. 1 следует, что TPS является естественным кубическим сплайном с двумя звеньями $[\alpha, 0]$ и $[0, \beta]$. Его вид определяется параметрами γ , \mathbf{r}_0 и производными f'_α и f'_β в точках α и β . Зависимость от непрерывных параметров позволяет использовать TPS-модель в двух режимах: *статическом* (α , β фиксируются, τ изменяется) и *динамическом* (хотя бы один из параметров изменяется вместе с τ). Динамический режим используется на этапе слежения по кубическому участку кривой, а статический на стадии «соединения» сегментов в одну кусочно-кубическую аппроксиманту.

Если α и β ассоциировать с узлами a_k и b_k k -го отрезка, то «статическая» аппроксиманта C_k определится значениями коэффициентов $\mathbf{r}_{0k} = [r_{\alpha k}, r_{\beta k}, r_{0k}]^T$, θ_k и значениями параметров $\mathbf{A}_k = (a_k, b_k, x_{0k})$, $k = 1, 2, \dots$. На каждом из таких подынтервалов, при $r_{*k} \equiv f_{*k}$, уравнения (2)–(5) гарантируют равенство значений аппроксиманты и функции в узлах a_k , b_k и x_{0k} . Вне k -го отрезка считаем $C_k = 0$.

Величины θ_k , a_k и b_k нам неизвестны. Так как для кусочно-кубической аппроксиманты $b_k \equiv a_{k+1}$, то в режиме слежения достаточно находить только правую границу отрезка (b_k), которую будем называть «точкой останова».

2. СВЯЗ ПАРАМЕТРА θ С ПРОИЗВОДНОЙ В УЗЛАХ

В случае кусочно-полиномиальной интерполяции C^1 -гладкость обеспечивается равенствами значений функции и первых производных в узлах с соответствующими значениями приближающей функции. Прямое четырехточечное преобразование для f позволяет найти оценку значений производных в узловых точках f'_α и f'_β . По определению, операция ДРТ задается формулой

$$f_\tau^\triangleleft(\tau; \mathbf{a}, \mathbf{r}_0) \triangleq (\mathbf{r}, \mathbf{p}),$$

где $\mathbf{p} = [p_1, p_2, p_3]^T$ — вектор весовых функций, определяемых четверкой $\{0\tau\alpha\beta\}$. Функции p_i выражаются через d_i в виде $\mathbf{p} = [-d_1/d_3, -d_2/d_3, 1/d_3]^T$. Применение ДРТ-операции к модели (2) дает уравнение прямой относительно τ :

$$C^\triangleleft(\tau; \mathbf{a}, \mathbf{r}_0) = \Pi^\triangleleft(\tau; \mathbf{r}_0) + \theta Q^\triangleleft = f_0 + \theta\alpha\beta\tau,$$

так как $\Pi^{\triangleleft} = r_{\beta}p_1 + r_{\alpha}p_2 + \Pi_{\tau}p_3 = f_0$, а $Q^{\triangleleft} = Q_{\alpha}p_1 + Q_{\beta}p_2 + Q_{\tau}p_3 = 0p_1 + 0p_2 + Q_{\tau}p_3 = \alpha\beta\tau$.

В [10] получены формулы для вычисления f'_{α} и f'_{β} через $f_{\alpha}^{\triangleleft}$, $f_{\beta}^{\triangleleft}$, \mathbf{r}_0 и \mathbf{a} :

$$f'_{\alpha} = \frac{1}{\alpha\beta\gamma}[\alpha^2\Delta_{\beta\alpha} + \gamma^2(f_{\alpha} - f_{\alpha}^{\triangleleft})], \quad (7)$$

$$f'_{\beta} = \frac{1}{\alpha\beta\gamma}[\beta^2\Delta_{\beta\alpha} - \gamma^2(f_{\beta} - f_{\beta}^{\triangleleft})], \quad (8)$$

где $\gamma = \beta - \alpha$, $\alpha \neq \beta$; $\alpha, \beta \neq 0$, а $\Delta_{\beta\alpha} = f_{\beta} - f_{\alpha}$, $f_{*}^{\triangleleft} = f^{\triangleleft}(\tau_{*})$.

Заменяя $f_{\alpha}^{\triangleleft}$ и $f_{\beta}^{\triangleleft}$ в (7) и (8) на $C_{\alpha}^{\triangleleft}$ и $C_{\beta}^{\triangleleft}$ соответственно, найдем выражения оценок производных в узловых точках через α , β , \mathbf{r}_0 и θ :

$$\hat{f}'_{\alpha} = \frac{1}{\alpha\beta\gamma}[\alpha^2\Delta_{\beta\alpha} + \gamma^2(\Delta_{\alpha_0} - \theta\alpha^2\beta)], \quad (9)$$

$$\hat{f}'_{\beta} = \frac{1}{\alpha\beta\gamma}[\beta^2\Delta_{\beta\alpha} - \gamma^2(\Delta_{\beta_0} - \theta\alpha\beta^2)]. \quad (10)$$

Из (9) и (10) легко получить формулу, выражающую $\hat{\theta}$ через γ , \mathbf{r}_0 и оценки производных \hat{f}'_{α} , \hat{f}'_{β} на концах сегмента:

$$\hat{\theta} = \frac{1}{\gamma^2} \left(\hat{f}'_{\alpha} + \hat{f}'_{\beta} - \frac{2}{\gamma}\Delta_{\beta\alpha} \right). \quad (11)$$

Формула (11) устанавливает связь коэффициента θ с основными параметрами локального кубического сплайна (2), включая длину отрезка γ . Поэтому C^1 -гладкость *существенно* зависит от точности $\hat{\theta}$ и \mathbf{r}_0 на каждом звене.

Если узлы α и β известны, а \mathbf{r}_0 , \hat{f}'_{α} и \hat{f}'_{β} заданы с хорошей точностью, то формулы (2)–(5) и (11) позволяют построить кусочно-кубическую аппроксиманту, т. е. решить задачу в постановке эрмитовой кубической интерполяции [3].

Оценки \hat{f}'_{α} и \hat{f}'_{β} можно найти по (9) и (10). Если вектор \mathbf{r}_0 задан, то для решения задачи локальной кубической аппроксимации достаточно найти $\hat{\theta}_k$ и концы промежутка a_k и b_k . Оценку $\hat{\theta}_k$ можно найти либо рекурсивным методом наименьших квадратов (РНК) (процедура типа Роббинса–Монро) [11], либо последовательным усреднением точечных оценок $\hat{\theta}_m$, определяемых по специальным четырехточечным наборам. Значение b_k определяется *динамически*, в процессе рекуррентного вычисления оценки $\hat{\theta}_k$, которая должна сохранять постоянное значение на кубическом участке кривой (критерий «точки останова»).

3. ПОИСК «ТОЧКИ ОСТАНОВА»

Итак, нам надо найти b_k и $\hat{\theta}_k$ по входным данным. Если $f(\tau) \approx C(\tau; \cdot)$, $\tau \in [\alpha, \beta]$ и $\alpha < 0 < \tau_* < \beta$, то при малой дисперсии входных ошибок точечную оценку $\hat{\theta}_*$ для соответствующей четверки можно найти из уравнения модели (2) в виде

$$\hat{\theta}_* = (f_* - \Pi_*)/Q_*. \quad (12)$$

Рассмотрим более подробно алгоритм поиска координат узловых точек на кривой f . Из двух фиксированных стартовых точек (x_α, r_α) , (x_0, r_0) и пары текущих (x_m, f_m) и (x_{m+s}, f_{m+s}) сформируем четверки (рис. 2) и по (12) вычислим $\hat{\theta}_m$ на участке, где $f \approx C$; значения $\hat{\theta}_m$ должны оставаться *почти* постоянными. Здесь через m обозначен индекс точки в массиве входных данных, s равно числу «пропущенных» точек между τ_m и $\beta_n \equiv \tau_{m+s}$. Буквой n обозначается индекс четверки на текущем отрезке.

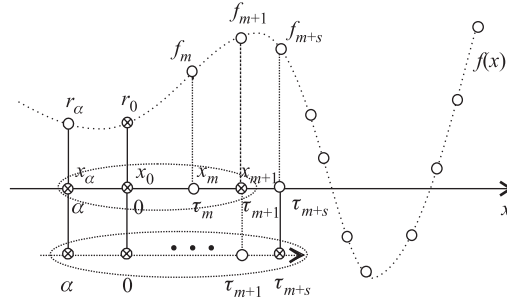


Рис. 2. Схема выбора четверок

Необходимо обратить внимание на правило выбора параметров α и x_0 , т. е. стартовых точек. Так как τ_m и β_n при слежении удаляются от нулевой точки, то для увеличения базы $\beta_n - \alpha$ точку α надо брать слева от нуля ($\alpha < 0$).

Абсциссы крайних точек в таких четверках ассоциируются с параметрами α и $\beta_n \equiv \tau_{m+s}$, нуль — с базисной, а τ_m — с текущей (пробной) точкой. После подстановки этих параметров в (3) и (4) найдем $\hat{\theta}_m$ по формуле (12) и последовательно вычислим усредненную оценку $\bar{\theta}_n$ по следующей рекуррентной формуле:

$$\bar{\theta}_n = [(n-1)/n]\bar{\theta}_{n-1} + \hat{\theta}_m/n; \quad n = 1, 2, \dots \quad (13)$$

Другой способ вычисления оценки θ_n приводится ниже (см. формулу (21)).

На каждом шаге используем $\bar{\theta}_n$ для вычисления невязки

$$\delta_m = f_m - \Pi(\tau_m; \mathbf{a}_n, \mathbf{r}_{0n}) - \bar{\theta}_n Q(\tau_m, \alpha, \beta_n), \quad (14)$$

где $\mathbf{a}_n = (\alpha, \beta_n, x_0)$, а $\mathbf{r}_{0n} = [r_\alpha, r_0, f_{\beta_n}]^T$ — репер для четверки с индексом n .

Конец сегмента или момент «точки останова» определяется неравенством

$$|\delta_m| > T_\delta, \quad (15)$$

где T_δ — заданный уровень точности (см. разд. 5). Значения координат (τ_m, f_m) в последней четверке, прошедшей тест (15), принимаются за правый конец текущего промежутка (b_k, f_{b_k}) и левый следующего $(a_{k+1}, f_{a_{k+1}})$, а значение $\bar{\theta} \equiv \bar{\theta}_n$ используется для вычисления \hat{f}'_{a_k} и \hat{f}'_{b_k} . На основе (12)–(15) легко получить адаптивный алгоритм с изменяющимися во времени параметрами модели.

При нормальном распределении $\hat{\theta}_m$ оценка $\bar{\theta}$ совпадает со средним в выборке объема n и является оптимальной и несмещенной, т. е. при небольшой дисперсии входных ошибок формулы (12) и (13) гарантируют точность вычисления \hat{f}'_{a_k} , \hat{f}'_{b_k} и $\hat{\theta}$ по формулам (9)–(11). Ошибка вычисления $\hat{\theta}_m$ регулируется параметром s и имеет порядок (n^{-2}) по отношению к максимальной входной ошибке на промежутке (см. (17)).

В момент обнаружения «точки останова» координаты точек (a_k, f_{a_k}) и (b_k, f_{b_k}) в предпоследней четверке принимаются за узлы k -го сегмента. Для более гладкой стыковки со следующим звеном его начало надо временно ассоциировать с точкой $(\tau_{m-1}, f_{\tau_{m-1}})$. Отметим, что в процедуре поиска узла значение β изменяется от точки к точке, а значения α и x_0 фиксируются и, в определенном смысле, обеспечивают память алгоритма на локальном отрезке.

После обнаружения актуальных границ отрезка формируются его «статические узловые» параметры $\mathbf{A}_k = (a_k, b_k, x_{0k} \approx (a_k + b_k)/2)$ и соответствующий им вектор \mathbf{R}_0 . Значение x_{0k} , по возможности, выбирается в середине отрезка.

4. АЛГОРИТМ LOCUSD

По приведенным выше формулам построим алгоритм поиска локальных кубических сегментов по входным данным (Local Cubic Segments Detector — LOCUSD). При поиске k -го сегмента в этом алгоритме выделяются два этапа.

На *первом* этапе используется TPS-модель в *динамическом* режиме для слежения по кубическому участку кривой, после чего находятся $\bar{\theta}_k$, узлы \mathbf{A}_k и реперные ординаты \mathbf{R}_{0k} . На *втором* этапе корректируется уровень C^1 -гладкости в узлах. Для этого используются полученные на первом этапе

\mathbf{A}_k , \mathbf{R}_{0k} и $\bar{\theta}_k$ для вычисления \hat{f}'_{a_k} , \hat{f}'_{b_k} по (9), (10) и уточнения $\hat{\theta}_k$ по (11), т. е.

$$\hat{\theta}_k = \frac{1}{\gamma_k^2} \left(\hat{f}'_{a_k} + \hat{f}'_{b_k} - \frac{2}{\gamma_k} \Delta_{b_k a_k} \right), \quad \gamma_k = b_k - a_k. \quad (16)$$

Как отмечалось выше, модель (2) гарантирует совпадение значений функции в узлах, а гладкость аппроксиманты существенно зависит от точности \mathbf{R}_{0k} . Качество стыковки можно улучшить, если процесс слежения для k -го сегмента начинать с предпоследней точки, включенной в $(k-1)$ -й сегмент (*внахлест*).

Таким образом, основной цикл алгоритма LOCUSD на k -м сегменте может быть реализован следующими шагами (m — индекс точки на входе, n — индекс пробной четверки, k — индекс сегмента):

- шаг 1: $n = 0$;
- шаг 2: $n \leftarrow n + 1$; выбрать четверку, сформировать \mathbf{r}_{0n} и вычислить d_{im} , Π_m , Q_m , θ_m и δ_m по формулам (3)–(5), (12), (14), $i = 1, 2, 3$;
- шаг 3: если $(|\delta_m| > T_\delta)$, то перейти на шаг 6; («детектор узла»)
- шаг 4: вычислить $\bar{\theta}_n$ по формуле (13); $m \leftarrow m + 1$;
- шаг 5: перейти на шаг 2 (за следующей точкой);
- шаг 6: «breakpoint»: сформировать \mathbf{A}_k , \mathbf{R}_{0k} и вычислить \hat{f}'_{a_k} , \hat{f}'_{b_k} и $\hat{\theta}_k$ (16);
- шаг 7: запомнить \mathbf{A}_k , \mathbf{R}_{0k} и коэффициенты (6); $k \leftarrow k + 1$; выход.

Этот алгоритм легко программируется и имеет простую структуру (рис. 3). Он имеет простую вычислительную схему, а скорость алгоритма определяется в основном сложностью входных данных (сложностью формы кривой).

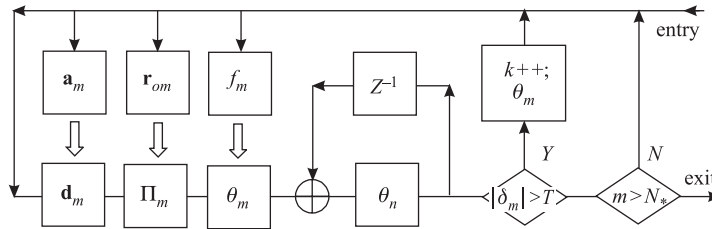


Рис. 3. Структура алгоритма LOCUSD

5. АНАЛИЗ УСТОЙЧИВОСТИ

Анализ устойчивости метода по отношению к погрешностям в исходных данных проведем на основе формул (9)–(12) и (13). Обозначим ошибки в значениях f_m через e_m , $m = 1, 2, \dots, N$. Чувствительность метода к e_m на

этапе слежения определяется ошибками в \mathbf{r}_0 и точностью вычисления оценки $\bar{\theta}_n$ по формуле (13). Ошибка $e_{\bar{\theta}}$ зависит от n и e_{θ_m} — ошибок вычисления $\hat{\theta}_m$ в точках τ_m по четверкам. Согласно (12) e_{θ_m} выражается через e_m и взвешенные ошибки $\mathbf{e}_0 = [e_\alpha, e_\beta, e_0]^T$ в виде

$$e_{\theta_m} = \frac{1}{Q_m} (e_m - d_{1m}e_\alpha - d_{2m}e_\beta - d_{3m}e_0).$$

С учетом выражений для Q_m и d_{im} в (3) и (5) получим

$$e_{\theta_m} = \frac{e_m}{\tau_m(\tau_m - \alpha)(\tau_m - \beta_n)} + \frac{e_\alpha}{\alpha\gamma(\tau_m - \alpha)} - \frac{e_\beta}{\beta_n\gamma(\tau_m - \beta_n)} - \frac{e_0}{\alpha\beta_n\tau_m},$$

т. е. при вычислении локальной оценки $\hat{\theta}_m$ ошибка e_{θ_m} преобразуется через α , β_n и τ_m в знаменателях. Поэтому вариации α и β_n влияют на e_{θ_m} . Например, с учетом размерности e_{θ_m} при $\alpha = -h$, $\beta_n = (n + s)h$, $\tau_m = nh$ и $s = 1$ получим

$$|e_{\theta_m}| < 4e_{\max}/[n(n + 2)], \quad n = 1, 2, 3, \dots, \quad (17)$$

где $e_{\max} = \max\{|e_m|, |e_\alpha|, |e_\beta|, |e_0|\}$, т. е. при вычислении параметра θ_m ошибки e_{θ_m} подавляются квадратично относительно n , а ошибка $e_{\bar{\theta}}$ убывает как (n^{-3}) , где n — индекс четверки на отдельном отрезке. Такой уровень подавления ошибок обеспечивает *устойчивость* процесса слежения и, в определенной степени, гарантирует точность вычисления \hat{f}'_α и \hat{f}'_β по формулам (9) и (10), влияющую на гладкость сопряжения звеньев. В формуле для e_{θ_m} значения α и β существенно влияют на точность оценки $\hat{\theta}_m$, т. е. они являются *параметрами аппроксимации (сглаживания)*. Дополнительный параметр $s_\beta \equiv s$ и аналогичный ему s_α служат для более тонкой настройки алгоритма по параметрам α и β .

Значения \hat{f}'_{a_k} и \hat{f}'_{b_k} вычисляются с использованием «узловых» параметров. С учетом e_0 и $e_{\bar{\theta}}$ ошибки вычисления производных в узлах преобразуются по формулам (9) и (10). Например, ошибка в значении производной в узле β равна

$$e'_\beta = -\frac{\beta}{\alpha\gamma}e_\alpha + \frac{2}{\beta}e_\beta - \frac{\gamma}{\alpha\beta}e_0 - \gamma\beta e_{\bar{\theta}}, \quad \text{где } \gamma \text{ — длина отрезка.}$$

Отсюда при $\alpha = -b$ и $b = \gamma/2$ получим

$$e'_b = (-e_a + 4e_b - 4e_0)/\gamma - \gamma^2 e_{\bar{\theta}}/2, \quad (18)$$

т. е. при вычислении оценок производных влияние ошибок реперных ординат на гладкость в точках сопряжения уменьшается с увеличением базы

$\gamma = b_k - a_k$. Так как $\gamma = nh$, то с учетом усреднения $\hat{\theta}_m$ и формулы (17) модуль ошибки $e_{\bar{\theta}}$ не превысит $4(n+2)e_{\max}/(n\gamma^3) \approx 4e_{\max}/\gamma^3$, где e_{\max} — максимальное значения модулей ошибок по всем точкам сегмента. Подставляя это выражение вместо $e_{\bar{\theta}}$ в (18), получим оценку максимальной ошибки вычисления f'_{b_k} по формуле (10)

$$|e'_b| < 11e_{\max}/\gamma. \quad (19)$$

Таким образом, ошибка вычисления оценки $\hat{\theta}_k$ по формуле (16) составит величину (γ^{-3}) по отношению к максимальной входной ошибке на сегменте. Полученные оценки должны учитываться при выборе значения T_{δ} .

Следует также отметить, что для сохранения устойчивости слежения при увеличении дисперсии входных ошибок «реперные ординаты» r_{α} , r_{β} и r_0 необходимо усреднять по нескольким ближайшим точкам.

6. СГЛАЖИВАНИЕ

Более подробно четырехточечный метод сглаживания кубическими сегментами в режиме адаптации изложен в работах [11, 12]. Ниже рассматривается модификация этого метода, использующая параметрическое вычисление весовых (базисных) функций. Концептуально метод сглаживания в режиме слежения полностью совпадает с изложенным выше методом аппроксимации. Основные отличия связаны с процедурой рекуррентного вычисления оценки $\hat{\theta}$ на первом этапе и вычислением оценок реперных ординат на втором.

В случае *сглаживания*, при большой дисперсии входных ошибок ($e = \tilde{f} - f$), оценка параметра θ определяется на основе критерия МНК

$$\sum_{i=1}^n \left(\tilde{f}_i - \Pi_i(\mathbf{a}, \bar{\mathbf{r}}_0) - \theta Q_i \right)^2 \rightarrow \min_{\theta} \quad (20)$$

с использованием приближенных значений реперных ординат $\bar{\mathbf{r}}_0$. Так как размер выборки n нам заранее не известен, а функционал (20) является линейным по отношению к искомому параметру, то для вычисления оценки $\hat{\theta}$ используется формула РНК в виде

$$\hat{\theta}_n = \hat{\theta}_{n-1} + K_n \left(\tilde{f}_n - \Pi_n - \hat{\theta}_{n-1} Q_n \right), \quad \hat{\theta}_0 = 0, \quad n = 1, 2, \dots, \quad (21)$$

где $K_n = Q_n / \sum_{j=1}^n Q_j^2$ — коэффициент усиления, подавляющий ошибку e_n . Поведение коэффициента K_n существенно зависит от выбора параметров α и β_n , т.е. $K_n = K_n(\alpha, \beta_n)$. Актуальное значение β_n связано с подвижной правой границей сегмента, а параметры α и x_0 фиксируются. Трансформация

реперных ошибок e_0 в (21) зависит от коэффициента K_n умноженного на веса d_{in} , $i = 1, 2, 3$. Подставляя Q из (3) в выражение для K_n при $\alpha = -1$ и $\beta_n = (n - 1)$, получим

$$K_n = \frac{n(n - \alpha)(n - \beta_n)}{\sum_j j^2(j - \alpha)^2(j - \beta_j)^2} = \frac{n(n + 1)}{\sum_j j^2(j + 1)^2} = \frac{15n}{3(n + 1)^4 - 5(n + 1)^2 + 2}.$$

При таком выборе параметра β_n весовые функции d_{in} соответственно равны:

$$d_{1n} = 1, \quad d_{2n} = \frac{n + 1}{n - 1} \quad \text{и} \quad d_{3n} = \frac{-(n + 1)}{n - 1}, \quad n > 1,$$

т. е. влияние d_{in} на ошибки e_0 несущественно, так как с ростом n их значения близки к ± 1 . На рис.4 приведен график K_n в логарифмическом масштабе по обеим осям для $n = 1-75$. Для сравнения приведены также графики для n^{-1} , n^{-2} и n^{-3} в том же масштабе. Мы видим, что кривые K_n и $|d_{in}K_n|$, $i = 1, 2$ практически совпадают и располагаются между n^{-3} и n^{-2} , причем ближе к n^{-3} . Таким образом, эффективность подавления ошибок e_n и e_0 в процедуре (21) близка к $O(n^{-3})$. Как и раньше, процесс (21) контролируется условием (15). В момент обнаружения «breakpoint», предпоследнее значение β_n определит правую границу интервала.

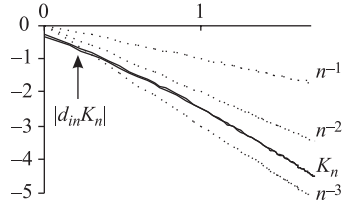


Рис. 4. $\log_{10} K_n$ и $\log_{10} |d_{in} K_n|$

На следующем этапе формируется $\mathbf{A}_k = (a_k, b_k, x_{0_k})$, где базисная точка x_{0_k} берется в середине отрезка $[a_k, b_k]$. После этого, используя найденные n и $\hat{\theta}$, по стандартной процедуре МНК находят оценки $\mathbf{R}_0 = [\hat{f}_{a_k}, \hat{f}_{b_k}, \hat{f}_{0_k}]^T$:

$$\sum_{i=1}^n \left(\tilde{f}_i - \Pi_i(\mathbf{A}, \mathbf{R}_0) - \hat{\theta} Q_i \right)^2 \rightarrow \min_{\mathbf{R}_0}.$$

Сглаживание отличается от аппроксимации способом вычисления оценки $\hat{\theta}$ и блоком вычисления $\hat{\mathbf{R}}_0$, поэтому шаги их алгоритмов практически совпадают.

7. ПРИМЕРЫ

Рабочие характеристики и эффективность алгоритма LOCUSD исследовались программно в среде Maple [13]. Алгоритм проверялся на моделированных и расчетных данных, а также на отсчетах реального сигнала от

радиоактивного источника, использующего зарядово-чувствительный преусилитель и стандартный спектрометрический усилитель [14].

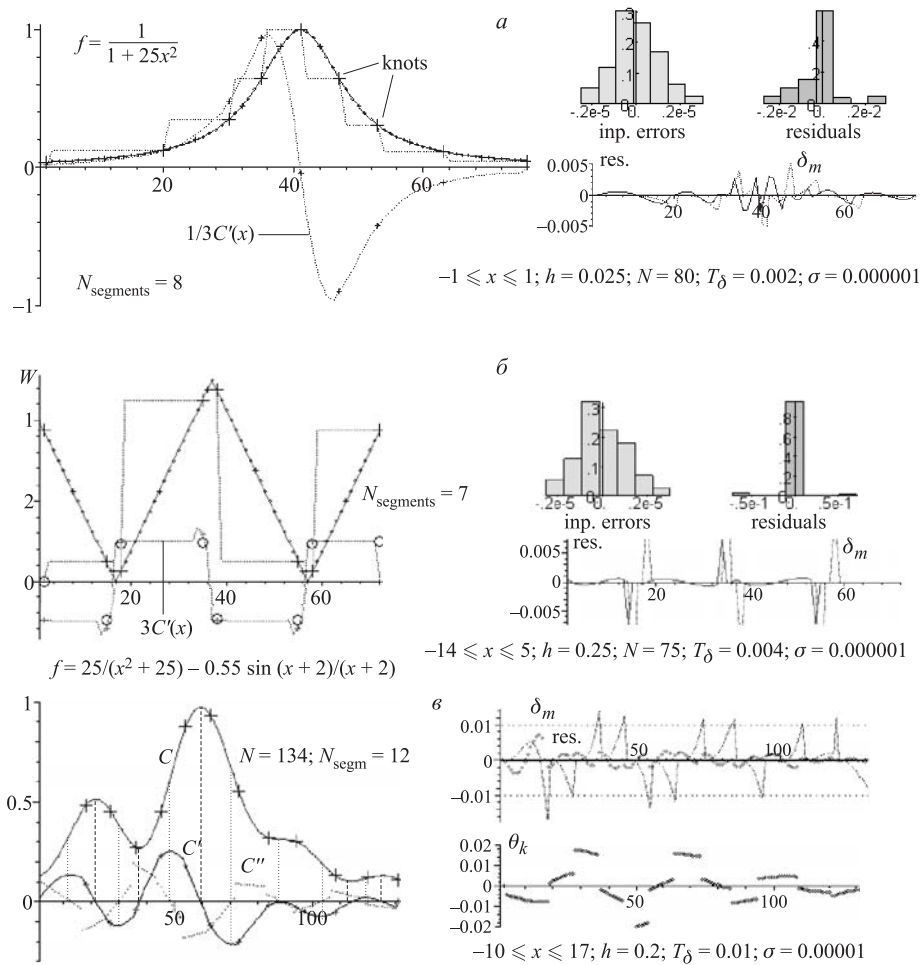


Рис. 5. Результаты поиска кубических аппроксимационных сегментов по входным данным в режиме автоматического слежения

На рис. 5, *a* приведен пример аппроксимации функции Рунге, заданной таблицей из 80 точек с постоянным шагом на отрезке $-1 \leq x \leq 1$. Входные данные изображены точками, а сплошной линией показана аппроксиманта.

Процесс слежения начинался фиксированием первых двух точек — $(\alpha = -h, f_\alpha)$ и (x_0, f_0) . Затем, в соответствии с шагами алгоритма LOCUSD, в режиме слежения находились узлы, отмеченные большим знаком «+». Малым «плюсом» отмечены точки (x_{0k}, f_{0k}) . Алгоритм обнаружил 8 локальных сплайнов. Найденные алгоритмом кубические сегменты отмечены пунктирной «лестницей». Приведен также график первой производной $C'(x; a_k, b_k, \mathbf{R}_{0k}, \hat{\theta}_k)$, полученной дифференцированием найденной аппроксиманты (чтобы сохранить масштаб, график производной сжат в три раза по оси ординат). Крестиками на кривой C' отмечены значения производной в узлах. Плавность сопряжения производной в узлах характеризует C^1 -гладкость. Из графика видно, что нуль производной соответствует точке максимума. Справа показаны гистограммы входных ошибок и остатков на выходе, а также приведены графики остатков и значений δ_m .

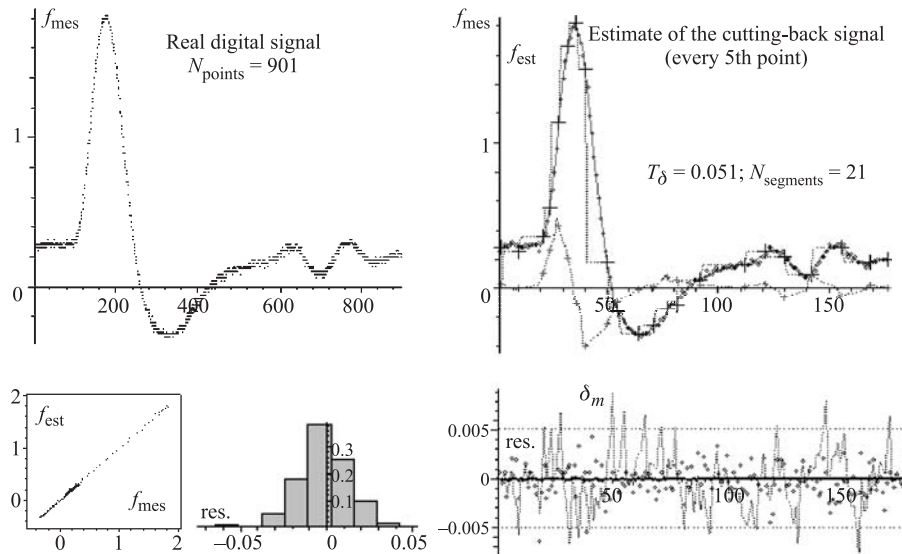


Рис. 6. Сглаживание реального цифрового сигнала алгоритмом LOCUSD

На рис. 5, б показана аппроксимация W-образной функции локальными кубическими сплайнами в режиме слежения. По графику производной аппроксиманты видны трудности приближения в точках изгиба.

Результаты кусочно-кубической аппроксимации функции, график которой содержит экстремумы и точки перегиба, показан на рис. 5, в. На вход алгоритма подавалась последовательность координат (x_m, f_m) , $m = 1, 2, \dots, 134$,

полученных на отрезке $-10 \leq x \leq 17$ с шагом $h = 0,2$ для функции

$$f(x) = 25/(x^2 + 25) - 0.55 \sin(x + 2)/(x + 2) + e(x),$$

где случайная ошибка $e(x) \sim N(0, \sigma^2)$ вычислялась для $\sigma = 0.00001$.

В результате, по 134 входным точкам найдено 12 кубических сегментов (сплошная тонкая линия), узлы которых показаны на графике. Внизу приведены графики первой и второй производных. Вертикальными пунктирными линиями отмечены экстремальные точки (минимумы, максимумы и точки

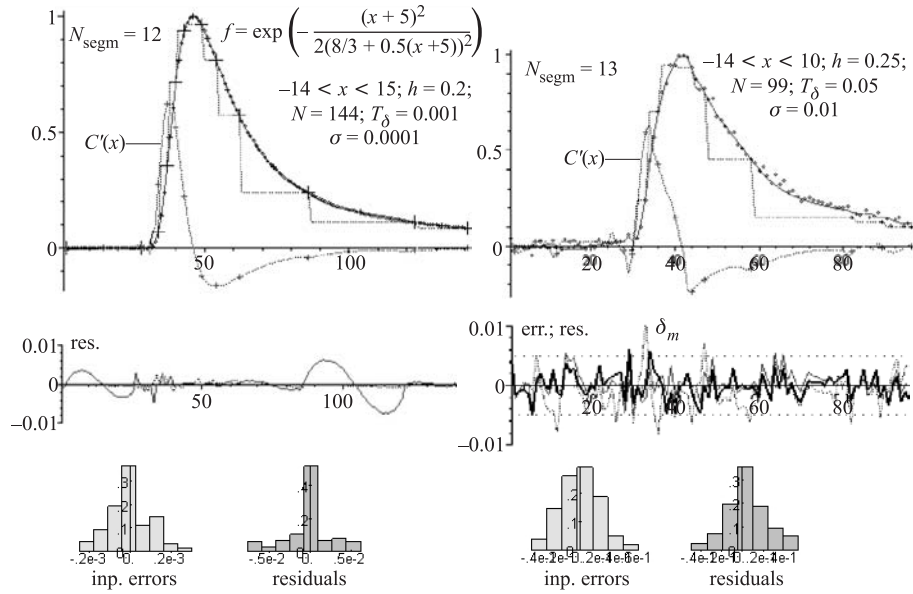


Рис. 7. Аппроксимация и сглаживание гауссоподобной функции

перегиба). Наблюдаются гладкость первой производной и точки разрыва второй производной. Справа приведены графики остатков res и δ_m («детектор узлов»). Внизу показаны графики изменений параметров θ_k в процессе пошагового усреднения.

На рис. 6 приводится пример обработки реального оцифрованного сигнала. Слева на рис. 6 показан 901 отсчет сигнала, полученного с осциллографа. Справа — результаты обработки прореженного сигнала (в обработку включалась каждая 5-я точка): узлы 21-го локального кубического сплайна, график глобальной оценки (сплошная линия), график производной с указа-

нием экстремальных точек (пунктирные линии). Внизу справа приведен график δ_m и график остатков. Слева внизу показана также гистограмма остатков.

Применение алгоритма LOCUSD для приближения гауссоподобной функции

$$f = A \exp\left(-\frac{(x - m)^2}{2(s + \omega(x - m))^2}\right)$$

демонстрирует рис. 7. Выбором параметров m , s и ω эта функция моделирует форму сигнала в радиационных детекторах. Показаны результаты аппроксимации сигнала для двух наборов данных, отличающихся только величиной дисперсии входных ошибок. На левом рисунке отмечены 12 локальных ку-

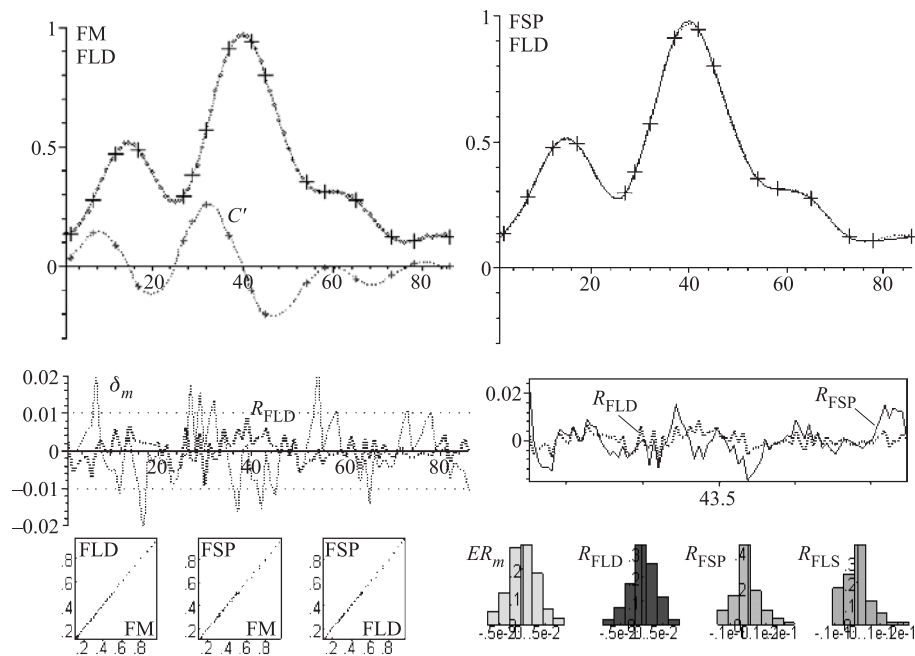


Рис. 8. Сравнение аппроксимант, полученных программой LOCUSD (слева) и процедурой Spline из пакета Maple (справа)

бических сегментов, обнаруженных программой по 144 точкам, приведены графики их производных, а также график распределения остатков на всем промежутке.

Справа приводятся результаты обработки аналогичного набора данных (99 точек, 13 локальных сегментов), с ошибками в сто раз большими, чем

в предыдущем наборе. Так как этот случай относится уже к *сглаживанию*, то шаг квантования был увеличен ($h = 0.25$), а реперные точки предварительно усреднялись по трем ближайшим. Точка экстремума, найденная после сглаживания, хорошо согласуется с реальной экстремальной точкой.

Полученные результаты подтверждают высокую устойчивость предложенного метода к входным ошибкам, а вид гистограммы остатков указывает на отсутствие смещений полученных точечных оценок.

На рис. 8 приведены результаты аппроксимации функции (рис. 5, *в*), полученные программой алгоритма слежения LOCUSD и стандартной процедурой *Curve Fitting [Spline] (XU,YU, x, degree=3)* из пакета Maple [13]. В результате анализа 89 точек ($-10 \leq x \leq 17$; $h = 0.3$; $\sigma = 0.003$; $T_\delta = 0.01$) программа нашла 17 узлов, по которым построила аппроксиманту FLD (график слева, пунктир). После этого по координатам обнаруженных узлов формировались векторы XU,YU и подавались на вход процедуры Spline. На ее выходе получен кубический сплайн FSP (график справа, сплошная линия). Внизу приведены графики и гистограммы невязок $R_{FLD} = FM-FLD$, $R_{FSP} = FM-FSP$ и $R_{FLS} = FLD-FSP$, а также гистограмма входных ошибок (FM – функция на входе). Сравнения обеих аппроксимант подтверждают работоспособность предложенного алгоритма.

ЗАКЛЮЧЕНИЕ

Предложенный метод кусочно-кубической аппроксимации и сглаживания таблично-заданных кривых в режиме автоматического слежения базируется на построении *параметрически-зависимых* весовых и базисных функций кубической модели. В рамках четырехточечной методологии [12], на основе модели трехточечного кубического сплайна (TPS) создан *простой в вычислении и устойчивый к ошибкам адаптивный алгоритм* для локальной аппроксимации и сглаживания кривых с *автоматическим обнаружением узловых точек*. Алгоритм имеет простую структуру, легко программируется, ориентирован на работу в режиме реального времени и может быть использован для разработки глобальных методов аппроксимации и сглаживания. Он не требует предварительного задания узловых точек, составления и решения системы линейных уравнений для нахождения кусочно-кубической аппроксиманты. Координаты узловых точек, полученные в процессе слежения, могут быть использованы для вычисления классического кубического сплайна или для построения адаптивных сеток в случаях аппроксимации функций многих переменных. Эффективность и работоспособность метода подтверждены примерами обработки модельных и реальных данных. На основе алгоритма LOCUSD разработаны MS Visual C# компоненты для кусочно-кубической аппроксимации в режиме автоматического слежения (APCA) [15]. Для вы-

полнения ARSA и конструирования аппроксимант на основе данных создано Windows-приложение, основанное на компонентах классов.

ЛИТЕРАТУРА

1. *Бартеньев О. В.* Фортран для профессионалов. Математическая библиотека IMSL. Ч.2. М.: Диалог-МИФИ, 2000.
2. *Калиткин Н. Н.* Численные методы. М.: ФМЛ, 1978.
3. *Каханер Д., Моулер К., Нэш С.* Численные методы и программное обеспечение. М.: Мир, 1998.
4. *Себер Дж.* Линейный регрессионный анализ. М.: Мир, 1980.
5. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов. М.: Мир, 1978.
6. *Сергиенко А. Б.* Цифровая обработка сигналов. СПб.: ПИТЕР, 2002.
7. *Friedman J.* A Variable span smoother. Department of Statistics Technical Report LCS5, Stanford University, Stanford, CA, 1984.
8. *Хардле В.* Прикладная непараметрическая регрессия. М.: Мир, 1993.
9. *Дикусар Н. Д.* Дискретные проективные преобразования на координатной плоскости // Математическое моделирование. 1991. Т. 10(3). С. 50–64.
10. *Dikoussar N. D.* // Comput. Phys. Commun. 1997. V. 99. P. 235–254.
11. *Дикусар Н. Д.* Кусочно-кубическое приближение и сглаживание кривых в режиме адаптации. Сообщение ОИЯИ Р10-99-168. Дубна, 1999.
12. *Дикусар Н. Д.* Методы 4-точечных преобразований в задачах аппроксимации и сглаживания кривых и поверхностей. Автореферат диссертации на соискание ученой степени доктора физико-математических наук. Дубна, 2002.
13. *Maple 9 Learning Guide.* Toronto: Maplesoft. Waterloo Maple Inc., 2003.
14. *Акимов Ю. К. и др.* Полупроводниковые детекторы в экспериментальной физике. М.: Энергоатомиздат, 1989.
15. *Торок Ч., Дикусар Н. Д.* MS.NET-компоненты для кусочно-кубической аппроксимации. Сообщ. ОИЯИ Р10-2004-202. Дубна, 2004.

Получено 1 декабря 2004 г.

Редактор *М. И. Зарубина*

Подписано в печать 27.01.2005.

Формат 60 × 90/16. Бумага офсетная. Печать офсетная.

Усл. печ. л. 1,12. Уч.-изд. л. 1,35. Тираж 310 экз. Заказ № 54760.

Издательский отдел Объединенного института ядерных исследований
141980, г. Дубна, Московская обл., ул. Жолио-Кюри, 6.

E-mail: publish@pds.jinr.ru

www.jinr.ru/publish/