РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ
№ 1 1993

УДК 681.3.058

Е. В. Рабинович, А. А. Рубан, М. П. Цапенко, Г. С. Шефель
(Новосибирск)

АЛГОРИТМЫ АДАПТИВНОЙ КУСОЧНО-ПОЛИНОМИАЛЬНОЙ АППРОКСИМАЦИИ

Рассмотрен ряд алгоритмов адаптивной кусочно-полиномиальной аппроксимации функций одной переменной, дается оценка коэффициентов сжатия данных для этих алгоритмов, приводятся результаты тестирования.

Для получения, передачи, обработки, хранения больших объемов измерительной информации возникает необходимость в применении методов, обеспечивающих компактное представление данных, при высоком быстродействии и заданной точности воспроизведения.

Приближение числовых данных удачно подобранным классом функций дает возможность представить эти данные значительно меньшим количеством параметров. Простота полиномиальной аппроксимации делает ее привлекательной для решения задач сжатия данных. Однако построение полиномов высоких степеней на больших интервалах определения данных сопряжено с рядом вычислительных трудностей. В такой ситуации представляется рациональным приближать входные данные на нескольких отрезках полиномами низких степеней. Для повышения эффективности сжатия данных отрезки приближения следует выбирать максимально допустимой длины. Такая возможность обеспечивается адаптивным построением аппроксимирующих полиномов. Полиномы строятся на отрезках, длина которых увеличивается до тех пор, пока ошибка приближения не превысит заданную допустимую погрешность.

В настоящей статье решается следующая задача. Пусть функция $f(x)$ задана своими значениями в $N + 1$ узлах отрезка $[a, b]$: $x_0, x_1, ..., x_N$, причем $x_i - x_{i-1} = h$, $i = 0, 1, ..., N$. Найти функцию $g(x)$, которая однозначно задается на отрезке $[a, b]$ с помощью M $(M < N)$ числовых параметров и приближает $f(x)$ со среднеквадратической погрешностью σ, не превосходящей наперед заданного фиксированного $\varepsilon > 0$. Характеристикой сжатия данных служит коэффициент сжатия $r(\varepsilon) = N/M$.

Будем искать $g(x)$ в виде кусочно-полиномиальной функции

$$g(x) = \{g_j(x) \in G(f),\ x \in [a_j, b_j),\quad j = 0, 1, ..., J\},$$

где $a = a_0 < b_0 < ... < b_J \le b < b_{J+1}$, $G(f)$ — класс полиномов степени m (в данной работе $m = 2, 3$), построенных по методу наименьших квадратов на отрезке $[a_j, b_j)$, содержащем максимальное количество $k + 1$ узлов, таких что

$$\sigma^2 = \sum_{i=0}^{k} [f(x_i) - g(x_i)]^2 < \varepsilon^2.$$

Будем считать, что для каждого отрезка аппроксимации выполнена замена переменных $t = (x - x_i)/h$, которая переводит его узлы в $0, 1, ..., k$, и вместо t снова будем писать x.

Полином m-й степени однозначно определяется $m + 1$ числовыми параметрами, например, коэффициентами при степенях переменной или отсчетами полинома в $m + 1$ фиксированных точках области определения. Для того чтобы восстановить $f(x)$ на отрезке $[0, k)$, необходимо знать эти параметры и количество узлов отрезка. Таким образом, коэффициент сжатия на данном отрезке $r_j(\varepsilon) = k/(m + 2)$. Как правило, числовые данные, являющиеся результатами измерений, представляются целыми числами невысокой разрядности, кроме того, наиболее экономное хранение и передача данных осуществляются на основе целых чисел. По этой причине, а также с целью повышения точности воспроизведения данных в качестве параметров аппроксимации, по-видимому, можно использовать не высокоразрядные действительные коэффициенты полинома, а его округленные значения.

Предлагаемые ниже алгоритмы оперируют классами полиномов, отличающимися способом построения многочленов и использованием условий сопряжения этих многочленов и их производных на краях отрезков аппроксимации.

Алгоритм 1. Адаптивная кусочная аппроксимация полиномами второй степени (ААП-2). Данный алгоритм использует класс $G(x)$ многочленов второй степени, заданных на конечном множестве узлов отрезка числовой прямой.

Полиномом наилучшего приближения второй степени на отрезке $[0, k]$ является многочлен, представляющий собой линейную комбинацию первых трех ортогональных многочленов Чебышева для k равноотстоящих узлов*:

$$g_j(x) = c_{j0}P_{k0}(x) + c_{j1}P_{k1}(x) + c_{j2}P_{k2}(x),$$

где $P_{k0}(x) = 1$, $P_{k1}(x) = 1 - 2x/k$, $P_{k2}(x) = 1 - 6x/(k - 1) + 6x^2/k(k - 1)$.

Коэффициенты c_{jn} вычисляются по формуле

$$c_{jn} = \frac{\sum_{i=0}^{k} f(i)P_{kn}(i)}{\sum_{i=0}^{k} P_{kn}^2(i)}.$$

Несмещенная оценка среднеквадратической погрешности приближения σ на j-м отрезке находится из равенства

$$\sigma_j^2 = \frac{\sum_{i=0}^{k} f^2(i) - \sum_{n=0}^{2} c_{jn}^2 \|P_{kn}(i)\|}{k - 2},$$

где $\|P_{kn}(i)\|$ — норма $P_{kn}(x)$.

Полином второй степени однозначно определяется тремя числовыми параметрами. Наиболее удобными для наших целей являются значения полинома $g_j(x)$ в первом, среднем и последнем узлах отрезка аппроксимации. Они вычисляются по формулам

$$g_j(0) = c_{j0} + c_{j1} + c_{j2},$$

$$g_j(k/2) = c_{j0} + (k + 2)c_{j2}/2(k - 1),$$

$$g_j(k) = c_{j0} - c_{j1} + c_{j2}.$$

Для восстановления функции $f(x)$ на отрезке $[0, k]$ используется формула

$$f(x) = g_j(0) - (3g_j(0) + g_j(k) - 4g_j(k/2))x/k + 2(g_j(0) + g_j(k) - 2g_j(k/2))(x/k)^2.$$

---

* Березин И. С., Жидков Н. Л. Методы вычислений.—М.: Физматгиз, 1962.—Т. 1.

Алгоритм аппроксимации заключается в следующем. В качестве трех параметров аппроксимации в память записываются значения $f(x)$ в первых трех узлах отрезка $[a, b]$. Это означает, что аппроксимации, как таковой, пока не произведено. Добавляется очередной узел, строится полином $g_j(x)$, приближающий $f(x)$ на отрезке $[0, 4]$, и вычисляются параметры $g_j(0)$, $g_j(k/2)$, $g_j(k)$. Если погрешность σ не превосходит ε, то эти параметры запоминаются вместо предыдущих. Количество узлов запоминается в качестве помеченного четвертого параметра аппроксимации, после чего строится полином для пяти узлов. Длина отрезка аппроксимации увеличивается до тех пор, пока σ не превысит ε. В этом случае формирование отрезка заканчивается, в памяти фиксируются параметры аппроксимации, полученные на предпоследней итерации. Следующий отрезок аппроксимации начинается с узла, отвергнутого на последней итерации. Процесс продолжается до полного исчерпывания узлов отрезка $[a, b]$. Если до конца отрезка остается один или два узла, то значения функции $f(x)$ в этих узлах записываются в память без обработки.

В результате, вместо последовательности значений $f(x)$ во всех $N + 1$ узлах $[a, b]$, получается массив данных, состоящих из значений $f(x)$ и/или $J + 1$ пачек чисел — трех значений $g_j(x)$, $j = 0, 1, ..., J$, и стоящего перед ними помеченного параметра, равного количеству узлов соответствующего отрезка аппроксимации.

Сжатие данных при таком порядке действий достигается, если функция $f(x)$ кусочно аппроксимируется полиномами второй степени на отрезках, среди которых есть хотя бы один отрезок, состоящий более чем из четырех узлов. Следовательно, для отрезка аппроксимации минимально разумной длины $[0, 3]$ нижняя граница коэффициента сжатия $\inf(r_j(\varepsilon)) = 1$. Дополнительное сжатие обеспечивается, если функция $f(x)$ аппроксимируется на некоторых отрезках полиномами нулевой или первой степени. В этих случаях вместо пачки из четырех чисел для аппроксимации необходимо два или три параметра.

Алгоритм 2. Адаптивная кусочная аппроксимация полиномами второй степени с учетом условий сопряжения (ААП-2С). Добиться уменьшения числа параметров, определяющих аппроксимирующую кусочно-полиномиальную функцию $g(x)$, можно, наложив на нее условие непрерывности или даже гладкости. При этом используется наличие смежных отрезков аппроксимации, имеющих общий граничный узел, и полиномы $gg_l(x)$ и $gg_{l+1}(x)$, $l = 0, 1, ..., L$, можно строить в классе $GG(x)$ многочленов, удовлетворяющих условию сопряжения:

$$gg_l(k_l) = gg_{l+1}(0),$$

где k_l — номер последнего узла l-го отрезка аппроксимации.

В этом классе полиномов на всех смежных отрезках, начиная со второго, один из пачки параметров аппроксимации — $gg_{l+1}(0)$ — уже определен. Таким образом, для двух смежных отрезков аппроксимации требуется семь параметров, для трех — десять и т. д. Следовательно, дополнительное сжатие данных обеспечивается даже для двух смежных отрезков, если их суммарная длина превышает семь узлов.

Соответствующий алгоритм отличается от предыдущего тем, что для аппроксимации функции $f(x)$ используются полиномы двух классов $G(x)$ и $GG(x)$: $g_j(x)$, $j = 0, 1, ..., J$, и $gg_l(x)$, $l = 0, 1, ..., L$.

Выполнение данного алгоритма начинается так же, как и предыдущего. Отличие возникает после построения полинома $g_j(x)$. Теперь осуществляется проверка: можно ли построить полином, приближающий $f(x)$ на отрезке, который образован последним узлом предыдущего отрезка аппроксимации и следующими тремя узлами. Если ответ положительный, то помечается, что $g_j(x) = gg_l(x)$, производятся построение полинома $gg_l(x)$ из класса $GG(x)$ и формирование пачки параметров аппроксимации, состоящей из трех чисел. По завершении построения полинома $gg_l(x)$ снова выполняется проверка. При отрицательном исходе проверки фиксируется факт $g_j(x) \in G(x)$ и процесс продолжается с начальной фазы — записи очередных трех значений функции $f(x)$ в память. Массив данных, полученный в результате работы данного алгоритма, состоит из значений $f(x)$ и/или $J+1$ пачек параметров аппроксимации полиномами класса $G(x)$, и/или $L+1$ пачек параметров аппроксимации полиномами класса $GG(x)$.

В данном случае на двух смежных отрезках, суммарная длина которых равна семи узлам, $\inf(r_l(\varepsilon)) = 1$.

Если в классе полиномов $GG(x)$ выделить подкласс $GD(x)$ полиномов, удовлетворяющих условию сопряжения производных в граничном узле смежных отрезков:

$$gg_l'(k_l) = gg_{l+1}'(0),$$

то возможно сокращение количества определяемых параметров аппроксимации еще на один, соответственно $\inf(r_l(\varepsilon)) = 1{,}167$. Этот параметр может быть вычислен, исходя из равенства

$$2c_{j2}k + c_{j1} = c_{j+1,1}.$$

К сожалению, выполнение условия сопряжения производных может вызвать появление эффекта «биений» (в случае когда у производных функций $f(x)$ и $gg_l(x)$ сильно различаются скорости изменения), что снижает качество аппроксимации. Принимая за оценку производной функции $f(x)$ первую разность $f(i+1) - f(i)$, можно осуществить слежение за разностью $f'(x) - gg_l'(x)$ с тем, чтобы вовремя оборвать процесс аппроксимации.

Другая возможность контроля связана с использованием идеи «щупа», заключающейся в следующем. При построении полинома $g_j(x)$ на отрезке $[0, k_j)$ на смежном отрезке $[k_j, 2k_j)$ строится вспомогательный полином $gh_j(x)$ («щуп»), такой что соблюдаются условия сопряжения:

$$g_j(k_j) = gh_j(k_j), \quad g_j'(k_j) = gh_j'(k_j),$$

среднеквадратическая погрешность аппроксимации на втором отрезке минимальна и среднеквадратическая погрешность аппроксимации на первом отрезке не превышает заданного ε. Максимальный k_j, для которого выполняются все эти требования, становится окончательной границей отрезка аппроксимации, причем гарантируется, что на смежном отрезке длиной не менее $k_j + 1$ существует полином $gg_l(x)$, приближающий функцию $f(x)$.

Алгоритм 3. Адаптивная кусочная аппроксимация трехзвенным квадратическим сплайном (ААС-2). Идея данного алгоритма является развитием темы сопряжения приближающих полиномов в граничных узлах отрезков с целью сокращения количества параметров аппроксимации.

Пусть $GS(x)$ — множество функций $s_j(x)$, $j = 0, 1, \ldots, J$, заданных на отрезке $[0, 3k_j]$, таких что на каждом из отрезков $[0, k_j]$, $[k_j, 2k_j]$, $[2k_j, 3k_j]$ эти функции являются квадратными трехчленами g_0, g_1, g_2, причем соответствующие параболы в узлах k_j и $2k_j$ удовлетворяют условиям гладкого сопряжения:

$$g_0(k_j) = g_1(k_j), \quad g_0'(k_j) = g_1'(k_j),$$

$$g_1(2k_j) = g_2(2k_j), \quad g_1'(2k_j) = g_2'(2k_j).$$

В силу данного определения функция $s_j(x)$ задается шестью независимыми параметрами, например: $g_0(0)$, $g_0(k_j/2)$, $g_1(3k_j/2)$, $g_2(5k_j/2)$, $g_2(3k_j)$ и число узлов k_j.

Построение полиномов второй степени производится так же, как и в алгоритмах 1 и 2, только с учетом условий гладкого сопряжения. Основная часть алгоритма заключается в построении функций $s_j(x)$, приближающих функцию $f(x)$ со среднеквадратической погрешностью, не превышающей σ на $J+1$

Функция	ААП-2			ААП-2С			ААС-2			ААС-3		
$\sin(ax^2)$	16	64	148	16	49	204	7	42	240	5	40	250
$e^{-bx}\cos(cx)$	13	52	191	15	46	220	7	42	240	5	40	250
dx	1	4	2500	1	4	2500	1	6	1667	1	8	1250

отрезках максимальной длины. Для данного алгоритма на минимальном отрезке $[0, 9]$ $\inf(r_j(\varepsilon)) = 1{,}667$.

Алгоритм 4. Адаптивная кусочная аппроксимация четырехзвенным кубическим сплайном (ААС-3). Отличие данного алгоритма от предыдущего заключается в построении полиномов третьей степени на четырех смежных отрезках аппроксимации. В условия гладкого сопряжения входят и вторые производные. Соответствующее количество параметров аппроксимации равно восьми. Для данного алгоритма на минимальном отрезке $[0, 13]$ $\inf(r_j(\varepsilon)) = 1{,}75$.

Результаты тестирования алгоритмов. В качестве тестовых использовались функции: $\sin(ax^2)$, $\sin(ax^3)$, $\exp(-bx)\cos(cx)$, $\exp(-ax)$, dx. Тестирование проводилось для значений $\sigma_1 = 0{,}03$ и $\sigma_2 = 0{,}05$. Для тестовых функций, имеющих колебательный характер, введен термин «горб», обозначающий отрезок функции от одного локального минимума (максимума) до другого. Параметры a и b подбирались такие, чтобы на 512 узлов приходилось 25—30 «горбов». Для затухающих тестовых функций параметры подбирались так, чтобы затухание происходило в e раз на 10—20 узлах. Таким способом достигался широкий диапазон изменения тестирующих функций как по частоте, так и по амплитуде.

Краткая сводка результатов тестирования с помощью наиболее выразительных функций приведена в таблице ($\sigma = 0{,}05$, $N = 9999$, $a = 0{,}0000005$, $b = 0{,}0003$, $c = 0{,}00005$, $d = 0{,}01$), где первое число обозначает количество отрезков аппроксимации, второе — количество параметров аппроксимации, третье — приближенный коэффициент сжатия.

Из таблицы результатов тестирования можно сделать выводы, что для принятых тестовых функций и исходных данных:

— введение условия сопряжения (ААП-2С) позволяет заметно увеличить коэффициент сжатия по сравнению с ААП-2;

— адаптивная аппроксимация на основе сплайнов (ААС-2, 3) приводит по сравнению с ААП-2, 2С к существенному уменьшению количества отрезков аппроксимации и к некоторому увеличению коэффициента сжатия.

Заметим, что на «хороших» участках тестовых функций (изменение частоты или амплитуды не более 20 %) коэффициенты сжатия имеют значительно более высокие значения.

К сожалению, отметим, что в алгоритме ААП-2С в варианте с сопряжением значений функции и производной, несмотря на принятые меры борьбы с «биениями», процесс аппроксимации вырождается.

Приведенные данные показывают, что рассмотренные относительно простые алгоритмы обеспечивают приближение функций на достаточно длинных отрезках аппроксимации с погрешностью, не превышающей заданную, и ощутимым коэффициентом сжатия.

Поступила в редакцию 23 ноября 1992 г.