УДК 621.391

Дмитриев Александр Георгиевич

КУСОЧНО-НЕПРЕРЫВНАЯ СЕГМЕНТАЦИЯ МНОГОМЕРНЫХ СИГНАЛОВ

Предлагается алгоритм кусочно-непрерывной аппроксимации структурных экспериментальных многомерных сигналов с заранее неизвестным числом интервалов разбиения на "однотипные" фрагменты. Построение многомерной кусочно-непрерывной аппроксимирующей функции выполняется "слева направо", что позволяет использовать метод динамического программирования для определения границ интервалов разбиения. Используется критерий качества аппроксимации, учитывающий количество данных на интервалах разбиения и "сложность" используемых локальных моделей сигналов.

Аппроксимация, сегментация, структурные сигналы.

Dmitriev Alexander Georgievich

ADAPTIVE ALGORITHM FOR COMPRESSION OF NOISY IMAGES

An algorithm for piecewise continuous approximation of structural experimental multidimensional signals with a previously unknown number of intervals for splitting signals into "similar" fragments is proposed. The construction of a multidimensional piecewise continuous approximating function is performed "left – to - right", which allows to use the dynamic programming method to determine the boundaries of the partition intervals. The approximation quality criterion is used, taking into account the number of data on the partition intervals and the "complexity" of the local signal models used.

Approximation, segmentation, structural signals.

Введение

В различных приложениях возникает проблема анализа так называемых структурных экспериментальных сигналов, рассматриваемых как последовательная во времени комбинация более простых сигналов (функций), которые обладают постоянными свойствами в соответствующие промежутки времени [1, 3, 5]. Обработка таких сигналов в большинстве случаев сводится к двухэтапной процедуре: выделению "однотипных" фрагментов (этап сегментации) и последующему построению описания представленных сигналов в целом. Использование существующих методов сегментации сигнала оказывается недостаточно эффективным в условиях высокой размерности, ограниченности экспериментальных наблюдений и неизвестном количестве "однотипных" фрагментов (интервалов разбиения). Кроме того, аппроксимирующая функция обычно имеет разрывы на границах "однотипных" фрагментов [2, 4].

Целью работы является разработка алгоритма кусочно-непрерывной аппроксимации многомерных сигналов с заранее неизвестным количеством интервалов разбиения сигналов на "однотипные" фрагменты, который при определенных условиях обеспечивает оптимальное значение выбранному критерию качества аппроксимации.

Постановка задачи

Пусть для анализа предъявлен набор s сигналов $y(t) = (y^{(1)}(t),...,y^{(s)}(t))$ (многомерный сигнал), в совокупности характеризующих исследуемый объект. Значения $y^{(i)}(t)$, $i = 1,...,s$, задаются в дискретные моменты времени $t = t_1,...,t_N$. Критерий качества аппроксимации J на выборке экспериментальных значений выбран в виде:

$$J = \frac{1}{N \cdot s}\sum_{i=1}^{s}\mu_i \sum_{j=1}^{r}\frac{n_j}{n_j - m}\sum_{t_l \in (T_{j-1},T_j]}[y^{(i)}(t_l) - F_j^{(i)}(t_l,\alpha_j^{(i)})]^2 \quad , \tag{1}$$

где μ_i - априорный «вес» сигнала $y^{(i)}$, n_j - число дискретных отсчетов на интервале $(T_{j-1},T_j]$, $F_j^{(i)}(t_l,\alpha_j^{(i)}) = \sum_{k=1}^{m}\alpha_{jk}^{(i)}\varphi_k(t)$ - многочлен по заданному набору базисных функций $\{\varphi_k(t), k = 1,...,m\}, \alpha_j^{(i)} = (\alpha_{j1}^{(i)},...,\alpha_{jm}^{(i)})$ - вектор оцениваемых параметров на j - ом интервале.

В критерии (1) используются веса $\mu_i, i = 1,...,s$, и $n_j / (n_j - m), j = 1,...,r$. Введение весов μ_i связано с тем, что на практике параметры $y^{(i)}$ часто имеют различную практическую значимость. Конкретные значения μ_i выбираются из содержательных соображений, обычно они нормируются так, чтобы $\sum_{i=1}^{s}\mu_i = 1$. Веса $n_j / (n_j - m)$ - это обычные нормирующие коэффициенты, учитывающие размерность модели.

Требуется найти такое разбиение $T = (T_0,T_1,...,T_r), T_0 < T_1 < ... < T_r$ (предполагается синхронное изменение сигналов $y^{(i)}$) заданного отрезка $[t_1,t_N]$, $T_0 = t_1, T_r = t_N$ на r интервалов $(T_{j-1},T_j], j = 1,...,r$ (r - в общем случае, неизвестно), и определить на каждом из этих интервалов такие значения векторов параметров $\alpha_j^{(i)}, i = 1,...,s$, чтобы функционал (1) принимал минимальное значение при условии ограничений на непрерывность аппроксимирующих функций:

$$F_j^{(i)}(T_j,\alpha_j^{(i)}) = F_{j+1}^{(i)}(T_j,\alpha_{j+1}^{(i)}), j = 1,...,r-1; i = 1,...,s. \tag{2}$$

Алгоритм

Обозначим через $\varepsilon^{(i)}(T_{j-1},T_j)$ - ошибку аппроксимации i - го сигнала на интервале $(T_{j-1},T_j]$, подсчитанную по методу наименьших квадратов. Тогда критерий (1) примет вид:

$$J = \frac{1}{N \cdot s}\sum_{i=1}^{s}\mu_i \sum_{j=1}^{r}\frac{n_j}{n_j - m}\varepsilon^{(i)}(T_{j-1},T_j) \quad . \tag{3}$$

Сделаем следующее предположение. Границы $T_j, j = 1,...,r-1$ разбиения $T = (T_0,T_1,...,T_r)$ и соответствующие локальные аппроксимирующие функции на каждом интервале разбиения будем находить «слева – направо». Пусть определено положение границы T_1, находим локальные аппроксимирующие функции $F_1^{(i)}(t,\alpha_1^{(i)})$, доставляющие минимум $\varepsilon^{(i)}(T_0,T_1)$, $i=1,...,s$. Далее фиксируем положение границы T_1, находим границу, доставляющую минимум $\varepsilon^{(i)}(T_1,T_2)$, $i=1,...,s$, при ограничении на непрерывность на границе T_1: $F_1^{(i)}(T_1,\alpha_1^{(i)})=F_2^{(i)}(T_1,\alpha_2^{(i)})$, $i=1,...,s$, фиксируем T_2 и т.д.

При этом предположении критерий (3) с ограничениями (2) обладает следующим свойством. Пусть T_j^* - некоторое фиксированное положение правой границы j - го интервала. Тогда границы $T_1^*,...,T_{j-1}^*$, полученные минимизацией (3) только по границам $T_1,...,T_{j-1}$, не зависят от значений границ $T_{j+1},...,T_{r-1}$. Действительно, функционал (3) можно представить в виде суммы двух неотрицательных величин: $J=J_a+J_b$, где

$$J_a=J_a(T_1,...T_{j-1}/T_j^*)=\frac{1}{N\cdot s}\left\{\sum_{k=1}^{j}\frac{n_k}{n_k-m}\sum_{i=1}^{S}\mu_i\varepsilon^{(i)}\left(T_{k-1},T_k\right)\right\},$$

$$J_b=J_b(T_{j+1},...T_{r-1}/T_j^*)=\frac{1}{N\cdot s}\left\{\sum_{k=j+1}^{r}\frac{n_k}{n_k-m}\sum_{i=1}^{S}\mu_i\varepsilon^{(i)}\left(T_{k-1},T_k\right)\right\},\ T_j=T_j^*.$$

Но тогда, очевидно, $\underset{T_1,...,T_{j-1}}{\arg\min} J=\underset{T_1,...,T_{j-1}}{\arg\min} J_a$. Из этого свойства следует, что если T_j^* - оптимальное положение j-ой границы, то и границы $T_1^*,...,T_{j-1}^*$, полученные минимизацией J по $T_1,...,T_{j-1}$, также оптимальны. Рассмотренное свойство критерия (3) позволяет воспользоваться процедурой динамического программирования [6] для определения оптимальных границ интервалов.

Пусть число интервалов равно r_0. Следующий рекуррентный алгоритм находит разбиение и локальные аппроксимирующие функции, доставляющие оптимальное значение функционалу (1) (при сделанном выше предположении).

Сначала последовательно табулируются функции $J_j(T_j)$, $j=1,2,...,r_0-1$, где

$$J_1(T_1)=\frac{1}{n_1\cdot s}\left\{\frac{n_1}{n_1-m}\sum_{i=1}^{S}\mu_i\varepsilon^{(i)}\left(T_0,T_1\right)\right\},\ T_1=t_p,...,t_{N-(r_0-1)p};\ J_j(T_j)=\frac{1}{s\cdot n(T_0,T_j)}\cdot$$

$$\cdot\min_{T_{j-1}=t_{(j-1)p},...,t_{T_j'-p}}\left\{s\cdot n\left(T_0,T_{j-1}\right)\cdot J_{j-1}\left(T_{j-1}\right)+\frac{n_j}{n_j-m}\sum_{i=1}^{S}\mu_i\varepsilon^{(i)}\left(T_{j-1},T_j\right)\right\},$$

$T_j = t_{j\cdot p},...,t_{N-(r_0-j)p}$; $j = 2,...,r_0 - 1$; T'_j-номер отсчета, соответствующий границе T_j, p - заданное минимально допустимое число отсчетов на интервале разбиения. Аппроксимирующие функции на соседних интервалах строятся с учетом условий на непрерывность (2). Одновременно запоминаются значения $M_{j-1}(T_j), T_j = t_{j\cdot p},...,t_{N-(r_0-j)p}$; $j = 2,...,r_0 - 1$, - значения оптимальных положений границ T_{j-1} для каждого T_j. Далее определяются оптимальные границы интервалов:

$$T^*_{r_0-1} = \underset{T_{r_0-1}=t_{(r_0-2)\cdot p},...,t_{N-p}}{\arg\min} \left\{ s \cdot n\left(T_0, T_{r_0-1}\right) J_{r_0-1}\left(T_{r_0-1}\right) + \frac{n_{r_0}}{n_{r_0} - m} \sum_{i=1}^{s} \mu_i \varepsilon^{(i)}\left(T_{r_0-1}, t_N\right) \right\},$$

$$T^*_{r_0-2} = M_{r_0-2}\left(T^*_{r_0-1}\right),...,T^*_1 = M_1\left(T^*_2\right). \tag{4}$$

Для нахождения разбиения $T^H = \left(T_0, T^H_1,...,T^H_{r_{H-1}}, T_r\right)$ при неизвестном числе интервалов используется экстремальный характер зависимости J от r. Действительно, при увеличении r, с одной стороны, происходит увеличение весов $n_j / (n_j - m + 1)$ - за счет уменьшения в среднем n_j, что, при прочих равных условиях, приводит к увеличению критерия (1). С другой стороны, при увеличении числа интервалов, происходит уменьшение обычной квадратичной невязки, что приводит к уменьшению (1). Одновременное действие этих факторов приводит к тому, что функционал (1) достигает своего минимального значения на некотором промежуточном (не граничном) значении r_H.

Для определения r_H можно воспользоваться подходом, предложенном в работе [2]. Сначала подсчитываются минимальные значения функционала (1) $J_j(t_N)$, $j = 2,...,r_{\max}$. Одновременно запоминаются значения $M_{j-1}(T_j)$ оптимальных положений границ T_{j-1} для каждого T_j. Далее производится выбор числа интервалов и определяются оптимальные границы. В качестве r_H выбирается такое наименьшее число интервалов r, при котором $J_r\left(t_N\right)$ - принимает минимальное значение. Оптимальные границы $T^H_1,...,T^H_{r_{H-1}}$ определяются из выражений, аналогичных (4):

$$T^H_{r_H-1} = M_{r_H-1}\left(t_N\right), T^H_{r_H-2} = M_{r_H-2}\left(T^H_{r_H-1}\right),...,T^H_1 = M_1\left(T^H_2\right).$$

Величина $r_{\max}$ выбирается из содержательных или статистических соображений. В частности, в качестве $r_{\max}$ можно использовать величину $[N / p]$, где $[x]$ - целая часть x.

Моделирование

Для проверки эффективности алгоритма было проведено его моделирование на специальных многомерных сигналах с заданными свойствами, генерируемых при помощи следующей процедуры.

Векторная функция рассматривается на интервале: $[1,N]$: $y(t)=\left(y^{(1)}(t),...,y^{(S)}(t)\right)$, где

$$y^{(i)}(t)=\sum_{j=1}^{r^*}\varepsilon_j\left\{\frac{y_j^{(i)}\cdot T_{j-1}-y_{j-1}^{(i)}\cdot T_j+\left(y_{j-1}^{(i)}-y_j^{(i)}\right)t}{T_{j-1}-T_j}\right\}+b\cdot\xi_t,\qquad (i=1,...,s) \tag{5}$$

представляет собой суперпозицию кусочно-линейной непрерывной функции и независимой гауссовой помехи с нулевым средним значением и дисперсией b^2; $T_0=1,T_{r^*}=N,$ $T_j, j=1,...,r^*-1-$ узловые точки кусочно-линейной векторной функции; $y_j^{(i)}$, $j=0,...,r^*-$ значения этих функций в точках T_j; ε_j-характеристическая функция равна единице, если $t\in\left(T_{j-1},T_j\right]$, и равна нулю, если $t\notin\left(T_{j-1},T_j\right]$. Узловые точки T_j, значения кусочно-линейной функции $y_j^{(i)}$ в этих точках, число интервалов r^* задаются с помощью генератора случайных чисел по следующему алгоритму:

1. $j=0;\ T_0=1;\ y_0^{(i)}=\varsigma_{0i},\ i=1,...,S.$
2. Если $N-T_j<p$, то переход к п. 4. В противном случае - переход к п. 3.
3. $j=j+1;\ T_j=T_{j-1}+\left[\beta_j c+p\right],\ y_j^{(i)}=\varsigma_{ji}\cdot d,\ i=1,...,s,$ переход к п. 2.
4. $T_j=N$; $r^*=j$.

ς_{ji},β_j-случайные числа равномерно распределенные, соответственно, на интервалах $[-1;1]$ и $[0;1]$; $p,d,c,N-$ заданные параметры алгоритма.

Далее находятся значения $y^{(i)}(t),\ (t=1,2,...,N),\ (i=1,...,s)$ вектор функции (5). В качестве локальных аппроксимирующих функций использовались линейные функции: $F_j^{(i)}\left(t,\alpha_j^{(i)}\right)=\alpha_{j1}^{(i)}+\alpha_{j2}^{(i)}\cdot t$, $i=1,...,s$, $j=1,...,r^*$.

В ходе исследований экспериментальный материал содержал три группы трехкомпонентных многомерных сигналов, полученных с использованием описанной выше процедуры. Первая группа состояла из многомерных сигналов без шума, вторая и третья, соответственно, со средним (или низким) и повышенным уровнем шума ($b^2\in[0,1-0,3]$). Как и ожидалось, при отсутствии шума алгоритм безошибочно находит требуемое число и границы интервалов аппроксимирующих функций.

Типичные зависимости критерия J от r для второй третьей групп сигналов показаны на рис. 1 и 2 (знак ↑ на рисунках указывает на фактическое количество интервалов).

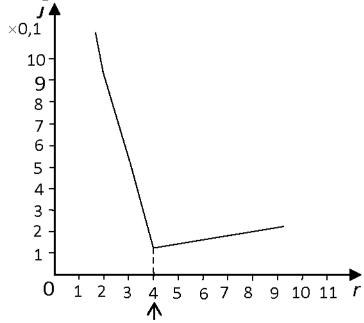

Рис. 1. Средний уровень шума

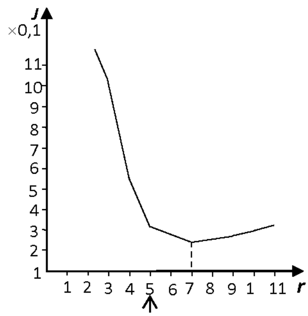

Рис. 2. Высокий уровень шума

Для сигналов со средним уровнем шума минимум функционала, как правило, приходится на искомое число интервалов. Для сигналов с повышенным уровнем шума несоответствие между оптимальным и искомым числом интервалов проявлялось чаще. Для обеих групп сигналов это расхождение наблюдалось, когда в многомерном сигнале были соседние интервалы, на которых разница в "поведении" сигнала была незначительной, и, как правило, такие интервалы содержали небольшое количество отсчётов.

Выводы

Таким образом, предложенный подход построения кусочно-непрерывной аппроксимирующей функции "слева направо" позволяет для определения границ интервалов разбиения применить метод динамического программирования. Для оценки количества интервалов разбиения используется экстремальное поведение критерия качества аппроксимации.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Моттль В.В., Мучник И.Б.* Скрытые марковские модели в структурном анализе сигналов. М.: Физматлит. 1999. 352с.

2. *Дмитриев А.Г.* Алгоритм оптимальной структурной аппроксимации экспериментальных многомерных сигналов. // Наукоёмкие технологии. 2010. №9. С. 31-35.

3. *Костин А. А. ,Красоткина О. В. , М. В. Марков М. В. , Моттль В.В., Мучник И.Б.* Алгоритмы динамического программирования для анализа нестационарных сигналов. // Журнал вычислительной математики и математической физики. 2004. Т. 44. №1. С. 70–86.

4. *Дмитриев А.Г., Дорофеюк А.А.* Методы кусочной аппроксимации многомерных кривых // Автоматика и телемеханика. 1984. № 12. - С.101-108.

5. *Браверман Э.М., Мучник И.Б.* Структурные методы обработки эмпирических данных. М.: Наука. 1983. 464с.

6. *Беллман Р., Дрейфус С.* Прикладные задачи динамического программирования. М.: Наука. 1969. 457с.

Дмитриев Александр Георгиевич, кандидат технических наук, доцент Военного университета радиоэлектроники, 162622, Россия, город Череповец, Советский проспект 126, телефон: +7 (921) 250-55-72, email: dag334a@fxmail.ru

Dmitriev Alexander Georgievich, Candidate of Technical Sciences (Ph.D.), associate Professor of the Military University of Radio Electronics, 162622, Russia, Череповец, Sovetsky Prospekt 126, phone: +7 (921) 250-55-72, email: dag334a@fxmail.ru