УДК 004.63

ИСПОЛЬЗОВАНИЕ СТОХАСТИЧЕСКОГО БАЗИСА В ЗАДАЧАХ ВОССТАНОВЛЕНИЯ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

В. И. Батищев, И. И. Волков, А. Г. Золин

Самарский государственный технический университет
443100, г. Самара, ул. Молодогвардейская, 244
E-mail: datatech@samgtu.ru

Рассмотрены вопросы построения базиса стохастических функций, представляющих собой выборки случайных чисел с заданным законом распределения. Представлен алгоритм построения аппроксимационной модели с использованием этого базиса. Дан общий алгоритм восстановления сигналов с помощью стохастического базиса методом наименьших квадратов с заданной весовой функцией. Предложен подход к решению задачи восстановления смазанных изображений при известной функции рассеяния точки путём построения модели обратного фильтра.

Ключевые слова: восстановление сигнала, реконструкция изображения, функция рассеяния точки, критерий адекватности, базисная функция, метод наименьших квадратов.

DOI: 10.15372/AUT20170415

Введение. В настоящее время теория аппроксимации нашла практическое приложение в различных областях науки и техники, таких как измерительная техника, диагностика, математическое моделирование сигналов, объектов и систем, теория идентификации и т. д. [1–4]. Одним из направлений, где аппроксимационный подход имеет перспективы, является восстановление сигнала по результатам регистрации откликов на выходе средств измерения, в частности при реконструкции изображений [5, 6]. Особенности формирования смазанных изображений в современных космических системах дистанционного зондирования Земли (КСДЗЗ), а также получение параметров функции рассеяния точки описаны в [7–9]. Достаточно распространённым является частный случай, когда функция рассеяния точки является функцией одной переменной, т. е. смаз происходит вдоль направления движения КСДЗЗ [10].

Каждая такая строка может быть представлена как свёртка значений строки исходного изображения x с известной функцией рассеяния точки h_0:

$$x_{\text{см}}(j) = \sum_{i=0}^{N_0-1} h_0(i) x(j-i), \tag{1}$$

где N_0 — величина смаза, $j = 0, \ldots, N-1$ (N — количество точек строки исходного изображения).

Постановка задачи. Задача реконструкции полученной с помощью (1) числовой последовательности сводится к нахождению функции h, представляющей собой весовую функцию обратного фильтра, позволяющего с использованием операции свёртки дать оценку значений пикселей восстановленного изображения $x_{\text{в}}$:

$$x_{\text{в}}(j) = \sum_{i=0}^{N-1} h(i)\, x_{\text{см}}(j-i).$$

Реконструкция полного изображения будет заключаться в применении данного алгоритма ко всем строкам исходного изображения.

В предлагаемой работе рассматривается подход, основанный на построении аппроксимационной модели искомого результата $x_{\text{в}}$. Несмотря на то что на данный момент теория аппроксимации разработана достаточно полно, при решении конкретных прикладных задач встают вопросы, связанные с обоснованием и выбором критерия соответствия модели оцениваемой функции (критерия адекватности), вида модели и системы базисных функций.

Как известно, системы базисных функций могут быть либо выбраны из множества известных, либо сформированы по определённой методике под решаемую задачу [11–14]. Решение проблемы формирования базисной системы, согласованной с априорной информацией и физической сущностью явлений, позволит повысить точность и достоверность результатов оценивания.

В данной работе рассматривается построение аппроксимационной модели в системе базисных функций, представляющих собой выборку случайных величин с заданным распределением, соответствующим распределению оцениваемой функции.

Построение аппроксимационной модели в базисе стохастических функций. Аппроксимация функции $x(n)$, $n = 0, \dots, N-1$, моделью

$$\hat{x}(n) = \sum_{k=0}^{p} C_k \varphi_k(n) \tag{2}$$

в базисе функций $\varphi_k(n)$ сводится к оцениванию коэффициентов C_k методом наименьших квадратов. Значение погрешности

$$\varepsilon = \sum_{n=0}^{N-1} (\hat{x}(n) - x(n))^2 \tag{3}$$

зависит от порядка модели p, значений параметров модели C_k и от вида базисных функций $\varphi_k(n)$, в качестве которых предлагается использовать сгенерированную выборку случайных величин с заданными параметрами распределения. Формально представляя в ортогональном базисе $\Psi_k(n)$ функции

$$\varphi_k(n) = \sum_{v=0}^{k} A_{k,\,v} \Psi_v(n) \tag{4}$$

и подставляя $\varphi_k(n)$ из (4) в (2), получим

$$\hat{x}(n) = \sum_{v=0}^{p} \beta_v \Psi_v(n),$$

где

$$\beta_v = \sum_{k=v}^{p} C_k A_{k,\,v}.$$

Из соотношения (3) находим, что погрешность будет минимальной, если

$$\beta_v = \frac{1}{\Phi_v} \sum_{n=0}^{N-1} x(n) \Psi_v(n).$$

Здесь

$$\Phi_v = \sum_{n=0}^{N-1} \Psi_v^2(n). \tag{5}$$

При этом минимальное значение погрешности $\varepsilon_{\min} = \sum\limits_{n=0}^{N-1} x(n)^2 - \sum\limits_{k=0}^{p} S_k$, где $S_k = \beta_k^2 \Phi_k$.

Нетрудно видеть, что для уменьшения погрешности $\varepsilon_{\min}$ k-ю базисную функцию следует выбирать такую, чтобы $S_k \to \max$.

На основании вышеизложенного и с использованием процедуры ортогонализации Грама — Шмидта получаем следующий алгоритм вычисления значений $A_{k,\,v}$ и β_k:

$$\Psi_0(n) = -1, \qquad \Phi_0 = N, \qquad A_{k,\,v} = \frac{1}{\Phi_v} \sum_{n=0}^{N-1} \varphi_k(n) \Psi_v(n), \quad v = \overline{0, k-1},$$

$$\Psi_k(n) = \varphi_k(n) - \sum_{v=0}^{k-1} A_{k,\,v} \Psi_v(n), \qquad \Phi_k = \sum_{n=0}^{N-1} \Psi_k^2(n), \qquad \beta_k = \frac{1}{\Phi_k} \sum_{n=0}^{N-1} x(n) \Psi_k(n).$$

Теперь для заданного k берётся набор выборок случайных чисел размером N с помощью программного генератора случайных чисел. В зависимости от задачи несложно сформировать выборки с различными законами распределения. Из этого набора выборок в качестве $\varphi_k(n)$ принимается та, при которой значение S_k максимально. Затем на основании соотношения (5) определяются коэффициенты

$$C_v = \beta_v - \sum_{k=v+1}^{p} A_{k,\,v} C_k, \qquad v = \overline{p, 0}, \tag{6}$$

которые применяются в (2) для построения аппроксимационной модели.

Восстановление сигнала с использованием стохастического базиса. Рассмотрим задачу восстановления сигнала (числового ряда), полученного из исходного по формуле (1) с известной весовой функцией h_0. Предлагаемый подход позволяет выполнять восстановление части числового ряда, состоящей из N последовательных элементов в диапазоне $[m - N + 1, \ldots, m]$. В качестве модели восстановленного числового ряда выберем следующую линейную зависимость:

$$\hat{x}_{\text{в}}(i) = \sum_{v=0}^{p} K_v \varphi_v(m - i). \tag{7}$$

Здесь p — порядок модели, m — номер последнего восстановленного значения части числового ряда, K_v — коэффициенты, $\varphi_v(i)$ — стохастические базисные функции, $i = 0, \ldots$ $\ldots, N + N_0 - 2$. Если эту модель подвергнуть воздействию того же прямого фильтра, что и исходный числовой ряд, то будем иметь

$$\hat{x}_{\text{см}}(j) = \sum_{i=0}^{N_0 - 1} h_0(i) \hat{x}_{\text{в}}(j - i)$$

или с учётом (7)

$$\hat{x}_{\text{см}}(j) = \sum_{v=0}^{p} K_v \theta_v(m - j), \tag{8}$$

где

$$\theta_v(j) = \sum_{i=0}^{N_0-1} h_0(i)\varphi_v(j+i), \qquad j = \overline{0, N_0+N-2}.$$

Значение параметров модели K_v будем вычислять, исходя из условия минимума квадратической погрешности с весовой функцией $w(j)$:

$$\varepsilon = \sum_{j=m-N+1}^{m} w(m-j)(\hat{x}_{\text{см}}(j)-x_{\text{см}}(j))^2 = \sum_{j=0}^{N-1} w(j)(\hat{x}_{\text{см}}(m-j)-x_{\text{см}}(m-j))^2, \quad N \geq p. \tag{9}$$

Здесь

$$\hat{x}_{\text{см}}(m-j) = \sum_{v=0}^{p} K_v\theta_v(j), \tag{10}$$

что следует из (8).

По аналогии с выражением (4) представим функции $\varphi_k(j)$ через некоторый ортогональный базис $\Psi_k(j)$ с весовой функцией $w(j)$:

$$\varphi_k(j) = \sum_{v=0}^{k} w(j)A_{k,v}\Psi_v(j).$$

Для этого будем использовать процедуру ортогонализации Грама — Шмидта:

$$\Psi_0(j) = \theta_0(j); \qquad \Phi_0 = \sum_{j=0}^{N-1} w(j)\Psi_0^2(j);$$

$$A_{v,k} = \frac{1}{\Phi_k}\sum_{v=0}^{N-1} w(j)\theta_v(j)\Psi_k(j); \qquad A_{v,v} = 1, \qquad k = \overline{0, v-1}, \quad v = \overline{1,p}; \tag{11}$$

$$\Psi_v(j) = \theta_v(j) - \sum_{k=0}^{v-1} A_{v,k}\Psi_k(j); \qquad \Phi_v = \sum_{j=0}^{N-1} w(j)\Psi_v^2(j).$$

Из (11) следует, что

$$\theta_v(j) = \sum_{k=0}^{v} A_{v,k}\Psi_k(j), \tag{12}$$

а также справедливо соотношение

$$\Psi_v(j) = \sum_{q=0}^{v} g_{v,q}\theta_q(j).$$

Функции $A_{v,k}$ и $g_{v,q}$ связаны соотношением

$$\sum_{v=k}^{q} g_{q,v}A_{v,k} = \sum_{v=k}^{q} g_{v,k}A_{q,v} = \delta_{k,q}. \tag{13}$$

Отсюда вытекает следующий алгоритм определения значений $g_{q,k}$:

$$g_{0,0} = 1; \qquad g_{q,q} = 1; \qquad g_{q,k} = -\sum_{v=k+1}^{q} g_{q,v}A_{v,k}, \qquad q = \overline{1,p}, \quad k = \overline{0, q-1}. \tag{14}$$

Для вычисления параметров модели подставим $\theta_v(j)$ из (12) в (10):

$$\hat{x}_{\text{см}}(m-j) = \sum_{k=0}^{p} \beta(k)\Psi_k(j),$$

где

$$\beta(k) = \sum_{v=k}^{p} A_{v,k} K_v. \tag{15}$$

Для обеспечения минимума взвешенной квадратической погрешности ε значения $\beta(k)$ определяются из выражения

$$\beta(k) = \frac{1}{\Phi_k} \sum_{j=0}^{N-1} \Psi_k(j) w(0) x_{\text{см}}(m-j). \tag{16}$$

Из (15) с учётом (13) получаем

$$K_v = \sum_{k=v}^{p} g_{k,v}\beta(k). \tag{17}$$

Подставив в (17) $\beta(k)$ из (16), будем иметь

$$K_v = \sum_{j=0}^{N-1} x_{\text{см}}(m-j) w(j) \sum_{k=v}^{p} \frac{g_{k,v}\Psi_k(j)}{\Phi_k}. \tag{18}$$

Приняв $i = m$ в выражении (7), запишем

$$\hat{x}_{\text{в}}(m) = \sum_{v=0}^{p} K_v \varphi_v(0).$$

И, наконец, с учётом (18) приходим к следующим алгоритмам нахождения значений весовой функции $h(j)$ обратного фильтра и восстановления числового ряда:

$$h(j) = w(j) \sum_{k=0}^{p} \frac{\Psi_k(j)}{\Phi_k} \sum_{v=0}^{k} g_{k,v}\varphi_v(0); \qquad \hat{x}_{\text{в}}(m) = \sum_{j=0}^{N-1} h(j) x_{\text{см}}(m-j). \tag{19}$$

Итак, предлагаемый алгоритм сводится к последовательному вычислению по формулам (8), (11), (14) и (19). Отметим, что значение весовой функции обратного фильтра $h(j)$ находится однократно, так как не зависит от ряда $x_{\text{см}}(j)$.

Восстановление смазанного изображения. При использовании алгоритма (19) для реконструкции смазанного изображения следует иметь в виду, что весовая функция $w(j) = 1$. Из соотношения (9) с учётом (16) определяем, что погрешность будет минимальной, если

$$\varepsilon_{\min} = \sum_{j=m-N+1}^{m} x_{\text{см}}(j)^2 - \sum_{v=0}^{p} S_v,$$

где $S_v = \beta(v)^2\Phi_v$. Таким образом, с помощью программного генератора для заданного v берётся набор выборок случайных чисел размером N. Из такого набора выборок в качестве $\varphi_v(i)$ принимается та, при которой значение S_v максимально.

Апробация алгоритма. Для данной работы были апробированы два алгоритма: построение аппроксимационной модели строки изображения и реконструкция смазанного эталонного изображения. В качестве эталонного взято изображение, полученное в результате дистанционного зондирования Земли.

Для исследования аппроксимационных свойств стохастических базисных функций по формуле (2) были построены модели с изменяемыми параметрами: порядком модели P и количеством точек оцениваемой функции N. Кроме этого, созданы аналогичные модели на основе полиномиальных и экспоненциальных базисных функций. По результатам моделирования была вычислена относительная среднеквадратическая погрешность

$$\Delta = \Big(\sum_{j=0}^{N-1}(x_{\text{в}}(j) - x(j))^2 \Big/ \sum_{j=0}^{N-1} x(j)^2\Big)^{1/2}, \tag{20}$$

показывающая степень отклонения модели от эталонного изображения. Данная величина может представлять собой численную оценку качества аппроксимации или реконструкции изображения.

В табл. 1 приведены значения погрешности аппроксимации в зависимости от выбранной системы базисных функций и параметров модели.

Для апробации алгоритма восстановления смазанного изображения выполнено моделирование смаза строк эталонного изображения по формуле (1) и восстановление по алгоритму (19). Это же изображение восстановлено с помощью других систем базисных функций:

Таблица 1

Погрешность аппроксимации в системе различных базисных функций

Количество точек	Порядок модели				
	10	15	20	25	30
Экспоненциальные базисные функции					
5	0,03225	0,02444	0,06422	0,08779	0,1028
9	0,003482	0,01142	0,03294	0,03743	0,0554
14	—	0,06213	0,05496	0,04668	0,04372
19	—	—	0,06162	0,06219	0,06626
29	—	—	—	—	0,02014
Полиномиальные базисные функции					
5	0,02572	0,02281	0,05506	0,07849	0,08912
9	$3{,}34 \cdot 10^{-7}$	0,02108	0,02671	0,03871	0,04504
14	—	0,04131	0,05635	0,06543	0,06292
19	—	—	0,04981	0,05763	0,05432
29	—	—	—	—	0,02067
Стохастические базисные функции					
5	0,005162	0,01605	0,05834	0,05735	0,08614
9	$1{,}786 \cdot 10^{-13}$	0,004244	0,008053	0,03076	0,02978
14	—	$3{,}484 \cdot 10^{-13}$	0,0004924	0,002921	0,01177
19	—	—	$8{,}478 \cdot 10^{-11}$	0,0002685	0,005382
29	—	—	—	—	$2{,}259 \cdot 10^{-13}$

Таблица 2

Погрешность восстановления изображения в различных базисах

Базис	Величина смаза			
	3	4	5	7
Без восстановления	0,0513	0,0704	0,0840	0,1164
Полиномиальный	0,0214	0,0246	0,0380	0,0412
Экспоненциальный	0,0258	0,0392	0,0489	0,0492
Стохастический	0,0192	0,0227	0,0362	0,0384

полиномиальных и экспоненциальных. Относительная среднеквадратическая погрешность восстановления Δ (20) представлена в табл. 2.

Несмотря на то что в настоящее время единые технические требования к точности реконструкции изображений для задач, решаемых КСДЗЗ, не определены, полученные результаты работы аглоритма можно косвенно оценить, сравнив их с погрешностью невосстановленного изображения. Эти данные приведены в первой строке табл. 2.

Заключение. В данной работе были получены и исследованы алгоритмы, использующие аппроксимационный подход для построения моделей и реконструкции смазанных изображений с помощью стохастических базисных функций. Вопрос применения полиномиальных, экспоненциальных и тригонометрических базисных функций к решению аналогичных задач восстановления сигналов и изображений рассматривался в работах [11, 12]. Приведённые результаты апробации (см. табл. 1 и 2) показывают, что стохастический базис позволяет уменьшить погрешность восстановления и аппроксимации, что даёт подходу, основанному на стохастических базисных функциях, преимущество при решении задач аппроксимации и восстановления данных. Выигрыш в точности в любом случае может быть обеспечен рациональным выбором вида базисных функций и метода оценивания параметров сформированных на их основе моделей. Выбор детерминированных базисных функций производится, как правило, субъективно на основании некоторых неформализованных предположений. В данном случае повышение точности обеспечивают предложенные алгоритмы генерирования последовательностей случайных чисел, образующих ортогональный базис, адаптируемый к статистическим свойствам обрабатываемых исходных данных в том смысле, что законы распределения сформированных последовательностей соответствуют закону распределения поступающих на обработку данных. При таком подходе выбор базисных функций и, следовательно, структуры модели осуществляется на основе объективных, доступных вычислению характеристик обрабатываемых данных.

СПИСОК ЛИТЕРАТУРЫ

1. **Алгоритмы** и программы восстановления зависимостей / Под ред. В. Н. Вапника. М.: Наука, 1984. 816 с.
2. **Батищев В. И., Мелентьев В. С.** Аппроксимационные методы и системы промышленных измерений, контроля, испытаний, диагностики. М.: Машиностроение-1, 2007. 393 с.
3. **Батищев В. И., Волков И. И., Золин А. Г.** Исследование аппроксимационных алгоритмов решения обратных задач технической диагностики // Вестн. СамГТУ. Сер. Технические науки. 2010. **28**, № 7. С. 6–13.
4. **Батищев В. И., Золин А. Г., Косарев Д. Н., Романеев А. Е.** Аппроксимационный подход к решению обратных задач анализа и интерпретации экспериментальных данных // Вестн. СамГТУ. Сер. Технические науки. 2006. № 40. С. 57–65.

5. **Батищев В. И., Волков И. И., Золин А. Г.** Исследование аппроксимационных свойств функциональных базисов в задачах реконструкции изображений при дистанционном зондировании земли // Тр. XVIII Междунар. конф. «Проблемы управления и моделирования в сложных системах» /Под ред. Е. А. Федосова, Н. А. Кузнецова, В. А. Виттиха. Самара: ООО «Офорт», 2016. С. 304–307.
6. **Клочко В. К., Кузнецов В. П.** Методы восстановления изображений и оценивания аппаратной функции по прореженной матрице наблюдений // Автометрия. 2016. **52**, № 6. С. 12–21.
7. **Злобин В. К., Еремеев В. В.** Обработка аэрокосмических изображений. М.: Физматлит, 2006. 288 с.
8. **Еремеев В. В., Егошкин Н. А.** Коррекция смаза изображений в системах космического наблюдения Земли // Цифровая обработка сигналов. 2010. № 4. С. 28–33.
9. **Кузнецов П. К., Семавин В. И., Солодуха А. А.** Алгоритм компенсации скорости смаза изображения подстилающей поверхности, получаемого при наблюдении Земли из космоса // Вестн. СамГТУ. Сер. Технические науки. 2005. № 37. С. 150–157.
10. **Кузнецов П. К., Мартемьянов Б. В.** Математическая модель формирования видеоданных, получаемых с использованием сканирующей съемки // Изв. Самар. науч. центра РАН. 2014. **16**, № 6-1. С. 292–299.
11. **Батищев В. И., Волков И. И., Золин А. Г.** Синтез цифровых КИХ-фильтров для решения задач восстановления сигналов с использованием критерия моментов // Вестн. СамГТУ. Сер. Технические науки. 2012. **36**, № 4. С. 98–105.
12. **Батищев В. И., Волков И. И., Золин А. Г.** Аппроксимационный подход к решению обратных задач восстановления сигналов в базисе экспоненциальных функций // Тр. XVI Междунар. конф. «Проблемы управления и моделирования в сложных системах» / Под ред.: Е. А. Федосова, Н. А. Кузнецова, В. А. Виттиха. Самара: Самар. науч. центр РАН, 2014. С. 678–681.
13. **Золин А. Г.** Исследование базисных систем на основе дробно-рациональных функций // Тр. V Всеросс. конф. с междунар. уч. «Математическое моделирование и краевые задачи». 2008. С. 55–58.
14. **Батищев В. И., Волков И. И., Золин А. Г.** Построение и оптимизация ортогональных базисных систем для аппроксимационного спектрально-корреляционного анализа и идентификации линейных динамических объектов // Вестн. СамГТУ. Сер. Технические науки. 2007. **20**, № 2. С. 47–52.

Поступила в редакцию 5 апреля 2017 г.