ПОСТРОЕНИЕ ПРОЕКЦИОННЫХ ОЦЕНОК ОДНОМЕРНЫХ ПЛОТНОСТЕЙ ВЕРОЯТНОСТЕЙ В ОРТОГОНАЛЬНЫХ БАЗИСАХ

Н.В. Аболмазова, Е.Е. Серповская. О.А Дегтярева
Самарский государственный аэрокосмический университет, г.Самара

Аналитическая обработка данных может производиться с применением аппроксимативных методов. Суть их заключается в нахождении подходящего аналитического выражения $\psi_k(x(t), \theta_0, \theta_1,...,\theta_n)$ с неизвестными параметрами $\theta_0, \theta_1,...,\theta_n$, удовлетворяющими заданному критерию оптимальности, которое бы описывало найденные экспериментальные результаты.

Задача аппроксимации (восстановления, оценивания) законов распределения (функций распределения, плотностей вероятности) на основе экспериментальных данных решалась многими авторами, разработавшими ряд методов ее решения, однако интерес к такого рода задачам сохраняется. Знание законов распределения случайной величины, случайного вектора, случайного процесса в аналитическом виде позволяет теоретическим путем определять законы распределения, характеристические функции, числовые характеристики других случайных объектов, связанных известным образом с первыми, определять вероятности одних событий через другие. Такие возможности позволяют значительно экономить время и средства, затрачиваемые на эксперимент.

Методы аппроксимации плотности вероятности можно разбить на три класса: методы параметрической аппроксимации, методы непараметрической статистики, методы многопараметрической статистики.

На практике нередко бывает затруднительно указать какие-либо веские причины, по которым конкретное распределение результатов наблюдений должно принадлежать тому или иному параметрическому семейству.

В этом случае целесообразно применение методов непараметрического оценивания: ядерное оценивание [1], метод вейвлет-анализа и др. Достоинством ядерных оценок плотности является их положительная определенность, что не всегда наблюдается у оценок, связанных с ортогональными разложениями. Однако последние имеют более лаконичную аналитическую форму. К многопараметрическим методам относятся гистограммный метод [2], аппроксимативный подход, например, сглаживание гистограммы с помощью сплайнов и линейно-независимых систем функций, в частности, системы ортогональных функций. К последним относятся гистограммно-аппроксимационный метод и метод проекционного оценивания, впервые предложенный Ченцовым Н.Н.

Проекционное оценивание плотности вероятности в ортогональных базисах.

Гистограммно-аппроксимационный метод [3] не всегда подходит к решению задачи. В ряде случаев можно использовать проекционное оценивание плотности вероятности.

В общем виде проекционная оценка плотности вероятности представляется частичной суммой ортогонального ряда:

$$f_n(x)=\sum_{i=0}^{k}\beta_{N,i}\psi_i(x)$$

Расчет коэффициентов разложения по базису $\beta_{N,i}$ производится следующим образом [4]:

$$\beta_{N,i}=\frac{1}{N}\sum_{j=1}^{N}\frac{\psi_i(x_j)}{\|\psi_i\|^2}$$

Для аппроксимации плотности вероятности, определенной на всей числовой оси, используем проекционную оценку с разбиением на ветви [5]:

$$f_n(x)=\sum_{i=0}^{k}(\beta_{N,i})_n\psi_i(x-x_0)1(x-x_0)+\sum_{i=0}^{k_1}(\beta_{N,i})_n\psi_i(x_0-x)1(x_0-x)$$

Методика расчета коэффициентов обеспечивающих гладкое склеивание, представленная в [5], является аналогичной методике, использующейся в гистограммно-аппроксимационном методе.

При использовании процедуры разбиения на ветви для вычисления среднеквадратической погрешности восстановления (аппроксимации) и теоретической погрешности следует использовать формулы такие же, как и для гистограммно-аппроксимационного метода [6].

Существует ряд универсальных систем, решающих задачу непараметрического оценивания плотности вероятности, которые отличаются широким спектром статистических методов, но не ориентированы на специфическую предметную область. Специализированные пакеты реализуют обычно методы, используемые в конкретной предметной области, доступ к ним весьма непрост.

Для исследования описанного метода аппроксимации плотности вероятности была спроектирована и разработана автоматизированная подсистема проекционного оценивания плотности вероятности, позволяющая получить проекционную оценку плотности вероятности в ортогональных базисах Лагерра, Лежандра и Дирихле. Данная подсистема входит в состав автоматизированной системы аппроксимации плотности вероятности в ортогональных базисах.

Остановимся подробнее на исследовании зависимости погрешности от масштабирующего коэффициента α при заданном числе слагаемых в аппроксимирующей сумме для проекционного метода, которое позволило провести разработанная система. Рассмотрим пример для закона распределения Рэлея в трех ортогональных базисах.

Изменяя k от 6 до 20 будем отмечать минимальные погрешности, а также интервалы для $\alpha_{лев}$ и $\alpha_{прав}$, в которых погрешность аппроксимации не больше 100% от минимальной. Как видно из результатов исследования (табл. 1) погрешность аппроксимации и теоретическая погрешность близки.

Таблица 1. Диапазоны $\alpha_{лев}$, $\alpha_{прав}$ для закона Рэлея

Значение k	Интервал для $\alpha_{лев}$	Интервал для $\alpha_{прав}$	Минимальная погрешность аппроксимации	Минимальная теоретическая погрешность
Для базиса Лагерра				
6	[2.4;5 2]	[2.4 6]	0,00550124 (3.2;3 2)	0.00282590 (3.8,2.6)
10	[0 5;2.3]	[0 5;5]	0,02054327 (1.3;0.9)	0,01449196 (1.5;0.9)
14	[0.2,1 6]	[0.2;5]	0,04316576 (0.8;0.6)	0,02957986 (0.8;0.4)
20	[0.1;0.6]	[0.1;5]	0,05471740 (0.35;0.35)	0,06073878 (0.8;0 35)
Для базиса Дирихле				
6	[0 3;3]	[0.4;3]	0,01325649 (0.5;0.8)	0,01528477 (0.5;0 7)
10	[0.05:0.35]	[0 05,0 25]	0,02719084 (0.1;0.15)	0,02796744 (0 1;0.15)
14	[2 2;7]	[3;7]	0,12647756 (5,4.6)	0,11613280 (5.4;4.2)
20	[3;6]	[4.75;7 25]	0,14155498 (6.25;4.75)	0,13192737 (6.25;4.5)
Для базиса Лежандра				
6	[0 39;3]	[0 24;0 66]	0.06249425 (0.8;0 38)	0.06220420 (2.06:0.38)
10	[0.25;5]	[0 05;1]	0,07285431 (0.8;0.15)	0,07623204 (0.7;0.15)
14	[0 25;4]	[0.2,1 5]	0,08363230 (0.75;0.25)	0,08047320 (0.5;0.25)
20	[0 25;4 25]	[0 2;2]	0,08456526 (1;0.25)	0,08212229 (0.5;0.25)

Отметим также следующие особенности:

1) с увеличением числа слагаемых в аппроксимирующей сумме для закона распределения Рэлея:

– для базиса Лагерра свойственно уменьшение значения минимального α;

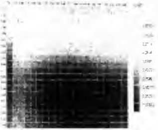
– для Дирихле свойственно увеличение;

– для Лежандра значения интервалов не изменяются;

2) для «поверхности» зависимости погрешности от α при заданном k свойственна «овраговая» структура:

– и в базисе Лагерра свойственна растянутость интервала для $\alpha_{прав}$ (см. рис. 1,а);

- в базисе Лежандра растянутость интервала для $\alpha_{псв}$ (см. рис. 1, б);
- в базисе Дирихле наблюдается отсутствие растянутости (см. рис. 1,в).



а) для базиса Лагерра

б) для базиса Дирихле

в) для базиса Лежандра

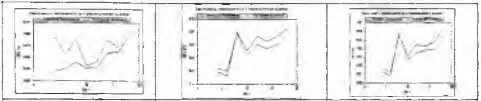
Рис. 1. «Поверхности» - зависимости погрешности от α при заданном k

На рис. 2 изображены графики исследования зависимости погрешностей от числа интервалов в аппроксимирующей сумме (k). Случайная выборка распределена по закону Рэлея, объем выборки меняется от 200 до 10000. Для аппроксимации применяются базисы Лагерра, Лежандра и Дирихле с использованием методики гладкого склеивания.

Основной вывод исследования: для фиксированного объема выборки существует число слагаемых в аппроксимирующей сумме, при котором теоретическая погрешность и погрешность аппроксимации теоретической плотности вероятности наименьшие. Ранее В.Н. Вапником [7] было отмечено, что неограниченное увеличение числа слагаемых в аппроксимирующей сумме не приводит к уточнению оценки, поскольку приближается выборка, которая сама случайна.

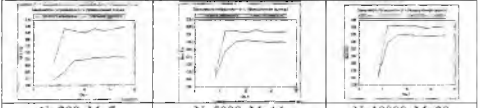
В работе были проведены подобные исследования для других одномодальных законов распределения, которые выявили аналогичные закономерности соотношения объема выборки и числа слагаемых.

Результаты исследований, полученные в ходе работы, предполагается дополнить анализом аппроксимации плотности вероятности проекционным методом многомодальных законов распределения, а также сравнительным анализом проекционного и гистограммно-аппроксимационного методов.



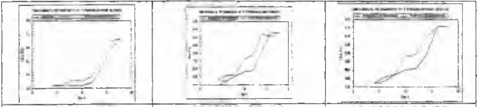
N=200, M=7	N=5000, M=14	N=10000, M=20

а) Графики погрешностей закона Рэлея в базисе Лагерра



N=200, M=7	N=5000, M=14	N=10000, M=20

б) Графики погрешностей для закона Релея в базисе Лежандра



N=200, M=7	N=5000, M=14	N=10000, M=20

в) Графики погрешностей закона Рэлея в базисе Дирихле

Рис. 2 Графики зависимости погрешностей от числа интервалов в аппроксимирующей сумме (k)

Список использованных источников

1.Коваленко, И.Н. Теория вероятностей и математическая статистика. Учебное пособие для втузов. – М.: Высшая школа, 1973. – 368 с.

2.Бочаров, П.П. Теория вероятностей. Математическая статистика: учебное пособие – М.: Гардарика. 1998. – 326 с.

3.Дегтярева О.А., Аболмазова Н.В., Серповская Е.Е. Восстановление плотности вероятности методом сглаживания гистограммы в ортогональных базисах. Сборник статей V Международной научно-практической конференции "Научное творчество XXI века". – Красноярск: Изд. Научно-инновационный центр, 2012. – С. 220-228.

4.Деврой Л., Дьерфи Л. Непараметрическое оценивание плотности. L1 – подход. Пер. с англ. – М.: Мир. 1988. – 408 с.

5.Дегтярева О.А. Аппроксимация гладких плотностей вероятности в ортогональных базисах, определенных на полуоси: Труды международного симпозиума "Надежность и качество–2007" – Пенза: ПГУ, 2007. – С. 180-184

6.Прохоров С.А., Дегтярева О.А. Автоматизированная система аппроксимативного анализа законов распределения случайных процессов. Вестник Самарского государственного технического университета. Выпуск 33. Серия "Технические науки" – Самара: СамГТУ. – 2005. – С. 335-340

7.Вапник В.Н. Восстановление зависимостей по эмпирическим данным. – М.: Наука, 1979. – 448 с.

ВЫБОР ОПТИМАЛЬНЫХ ПАРАМЕТРОВ ГАСИТЕЛЯ ПУЛЬСАЦИЙ ДАВЛЕНИЙ

П.В. Аболмасов

Самарский государственный аэрокосмический университет, г. Самара

Введение

Большое количество постановок технических или научных проблем можно сформулировать как задачу глобальной оптимизации (ГО). Не является исключением и задача выбора рациональных параметров гасителя пульсаций давления (ГПД) по критерию оценки среднего уровня акустической мощности, для решения которой можно использовать методы ГО целевой функции.

Описание физической модели гасителя пульсаций

Идея понижения акустической мощности шума от гасителя пульсаций давлений (см. рис. 1) заключается в установке специальных шайб. Тогда полную акустическую мощность, генерируемую гасителем пульсаций давления, можно рассчитать как сумму мощности клапана и мощностей каждой шайбы ГПД:

$$W = W_1 + \sum_{i=2}^{N} W_i. \qquad (1)$$

Рис. 1. Модель ГПД