

УДК 621.382.8

**ПРИМЕНЕНИЕ АСИМПТОТИЧЕСКИХ МЕТОДОВ
В АВТОМАТИЗИРОВАННОМ ПРОЕКТИРОВАНИИ
ТЕХНИЧЕСКИХ ОБЪЕКТОВ**

Анисимова Н.Т., Ивашов Е.Н., Князева М.П., Костомаров П.С.

*ФГАОУ ВПО «Московский институт электроники и математики
Национального исследовательского университета «Высшая школа экономики»»,
Москва, e-mail: eivashov@hse.ru; e-mail: ienmiem@mail.ru*

Рассмотрена проблема малых выборок, т.е. отсутствия необходимого количества эмпирических данных для принятия рационального конструктивно-технологического решения, при автоматизированном проектировании элементов, узлов и устройств технических объектов. Предложен рандомизированный подход определения коэффициентов размытости непараметрических решающих правил. Рассмотрено применение метода Лапласа в асимптотической оценке применяемых ядер в результате статистического моделирования.

Ключевые слова: автоматизированное проектирование; проблема малых выборок; непараметрические решающие правила; асимптотические методы.

**ASYMPTOTIC METHODS FOR COMPUTER-AIDED DESIGN
OF TECHNICAL OBJECTS**

Anisimova N.T., Ivashov E.N., Knyazeva M.P., Kostomarov P.S.

*FGAEU HPE «Moscow institute of electronics and mathematics The National research university
“High school of economics”», Moscow, e-mail:eivashov@hse.ru; e-mail: ienmiem@mail.ru*

The problem of small sample size, i.e. absence of required quantity of empirical data for decision making of rational constructive-technically, in case of computer-aided design of elements, nodes and technical objects' devices is considered. Randomization approach for blur factor's determination of nonparametric decision rules is suggested. The approach of Laplace's method in using cores' asymptotic bound in case of static model is considered.

Keywords: computer-aided design; small sample size problems; nonparametric decision rules; asymptotic methods.

При автоматизированном проектировании элементов, узлов и устройств технических объектов разработчики встречаются с проблемой малых выборок, т.е. отсутствия необходимого количества эмпирических данных для принятия рационального конструктивно-технологического решения [1].

Для «обхода» проблемы малых выборок при оценивании плотностей вероятностей $p(x)$ увеличим объем исходных данных $x^i, i = \overline{1, n}$, за счет результатов статического моделирования. С этой целью в β -окрестности каждой i -й точки выборки осуществим m имитаций с законом распределения $p_2(x)$. Полученная статическая выборка $x^i + x^j, j = \overline{1, m}, i = \overline{1, n}$, при равновероятных значениях $x^i, i = \overline{1, n}$, соответствует смеси плотностей вероятностей:

$$\hat{p}(x) = \frac{1}{n} \sum_{j=1}^m p_2^j(x_2). \quad (1)$$

Нетрудно заметить, что непараметрическая оценка (1) имеет вид

$$\bar{p}(x) = (nmc)^{-1} \sum_{i=1}^n \sum_{j=1}^m \Phi\left(\frac{x - x^i - x_2^j}{c}\right). \quad (2)$$

Существующий парадокс традиционных методов идентификации стохастических моделей состоит в сопоставлении конечной случайной выборки наблюдений переменных изучаемых объектов с конкретным набором параметров модели, оптимальным в некотором смысле. Предлагается рандомизированный подход определения коэффициентов размытости непараметрических решающих правил на основе процедуры их случайного выбора, который рассматривается на примере оптимизации алгоритмов формирования литографических процессов.

Впервые методика случайного выбора коэффициентов размытости ядерных функций при синтезе непараметрической оценки плотности вероятности была предложена в 1975 г. Т. Вагнером [1]. Формирование случайной последовательности коэффициентов размытости при оценивании плотности вероятности $p(x)$ осуществляется на основе выборки расстояний между исходными наблюдениями $(x^i, i = \overline{1, n})$ и их k -ближайшими соседями.

Рассмотрим **рандомизированный метод** оптимизации. Пусть $V = (x^i, i = \overline{1, n})$ – выборка из n статистически независимых

наблюдателей случайной величины $x \in R^1$ с плотностью вероятности $p(x)$, вид которой неизвестен. Будем считать, что $p(x)$ ограничена и непрерывна со всеми своими производными до второго порядка включительно. В качестве приближения по эмпирическим данным V искомой плотности вероятности $p(x)$ примем статистику типа Розенблатта – Парзена [2]

$$\bar{p}(x) = (nc)^{-1} \sum_{i=1}^n \Phi\left(\frac{x-x^i}{c}\right), \quad (3)$$

где $\Phi\left(\frac{x-x^i}{c}\right)$ – ядерные функции, удовлетворяющие условиям положительности, симметричности и нормированности; $c = c(n)$ – последовательность положительных чисел (коэффициентов размытости) таких, что

$$\lim_{n \rightarrow \infty} c(n) = 0, \quad \lim_{n \rightarrow \infty} nc(n) = \infty. \quad (4)$$

Свойства непараметрической оценки плотности вероятности (3), такие как асимптотическая несмещенность, состоятельность, сходимость почти наверное к $p(x)$, подробно рассмотрены в работе [3].

Для сравнения традиционного и рандомизированного метода оптимизации непараметрической оценки плотности вероятности определим отношение $\bar{W}(c^*) / \bar{W}_p(h^*)$ соответствующих им асимптотических выражений среднеквадратических критериев при оптимальных параметрах c и h .

Из условия минимума $W_p(c)$ и $\bar{W}_p(h)$ по c и h нетрудно получить

$$c^* = \left[\frac{\|\Phi\|^2}{n\|p^{(2)}(x)\|^2} \right]^{\frac{1}{5}}, \quad h^* = \left(\frac{5+t}{t} \right)^{\frac{1}{5}} c^*. \quad (5)$$

При оптимальных параметрах c^* , h^* отношение

$$\frac{\bar{W}(c^*)}{\bar{W}_p(h^*)} \sim \frac{(1+5t^{-1})^{\frac{1}{5}}}{1+t^{-1}} \quad (6)$$

меньше единицы при конкретных значениях параметра t закона распределения $p(c)$ коэффициентов размытия ядерных функций.

Однако использование непараметрической оценки со случайными значениями коэффициентов размытия ядерных функций

$$\tilde{p}(x) = n^{-1} \sum_{i=1}^n \frac{1}{c^i} \Phi\left(\frac{x-x^i}{c}\right) \quad (7)$$

позволяет снизить смещение при оценивании плотности вероятностей по сравнению с традиционной статистикой типа (3).

Можно показать, что асимптотическое выражение смещения

$$\bar{W}_p^1(h) =$$

$$= \int_0^h M(\bar{p}(x) - p(x)) p(c) \sim dc \frac{h^2(t+1)}{2(t+3)} p^{(2)}(x), \quad (8)$$

а его отношение к соответствующему смещению $\bar{W}^1(c)$ для традиционной непараметрической оценки при оптимальных параметрах c^* и h^*

$$\frac{\bar{W}_p^1(h^*)}{\bar{W}^1(c^*)} \sim \frac{t+1}{t+3} \left(\frac{5+t}{t} \right)^{\frac{2}{5}}. \quad (9)$$

Если параметр t плотности вероятности $p(c)$ больше или равен 2, то отношение (9) меньше единицы.

Анализ выражений (5), (8) показывает, что непараметрическая оценка плотности вероятности со случайными значениями коэффициентов размытости (7) обладает свойствами асимптотической несмещенности и состоятельности. Она характеризуется пониженным смещением (9) и несколько большим значением среднеквадратического отклонения (6) по сравнению с непараметрической статистикой (3). Следует ожидать проявления потенциальной эффективности непараметрической оценки плотности вероятности (7) при конечных объемах статистических данных.

Реализация «обхода» проблем малых выборок при оценивании плотностей вероятностей $p(x)$ обеспечивается также не только увеличением объема исходных данных x^i , $i = \overline{1, n}$, но и результатами технического моделирования, т.к. сложность и многообразие процессов функционирования проектируемых технических систем, таких как, к примеру, литографическое оборудование, не всегда позволяют получать для них адекватные математические модели, сформулированные в виде различных аналитических соотношений [4].

Рассмотрим применение метода Лапласа в асимптотической оценке применяемых ядер в результате статистического моделирования [5].

Методом Лапласа можно назвать ту совокупность приемов, способов оценок интегралов $\int_a^b K(t, \lambda) \varphi(t) dt$, когда с ростом параметра λ к положительной бесконечности ($\lambda \rightarrow \infty$), график по t ядра $K(t, \lambda)$ приобретает все более ярко выраженный вид профиля горной страны. Чем больше λ , тем выше (относительно) становятся «пики», глубже и шире (относительно) долины, круче склоны пиков. Не исключается случай, когда с изменением λ положение «пиков» меняется. К обобщениям метода Лапласа

можно отнести приемы оценки интегралов $\int_{\alpha(\lambda)}^{\beta(\lambda)} K(t, \lambda) \varphi(t, \lambda) dt$ с описанным выше характером ядра $K(t, \lambda)$. На рис. 1 представлено несколько примеров таких ядер [6].

$$K(t, \lambda) = e^{-\lambda t}, (0 \leq t < +\infty) \quad (10)$$

$$K(t, \lambda) = e^{-\lambda t^2}, (-\infty < t < +\infty) \quad (11)$$

$$K_n(t) = \left(\frac{1}{1+t^2} \right)^n. \quad (12)$$

Растет и количество сравниваемых по величине пиков; когда с ростом λ один или несколько пиков «наезжают» на особую точку функции $\varphi(t)$, или вообще любую комбинацию «неприятностей» подобного рода – особенно, если рассматривается интеграл вида $\int_{\alpha(\lambda)}^{\beta(\lambda)} K(t, \lambda) \varphi(t, \lambda) dt$. Поэтому основная трудность здесь не в доказательстве теоремы, а, пожалуй, в отборе тех результатов, которые следует явно сформулировать [6].

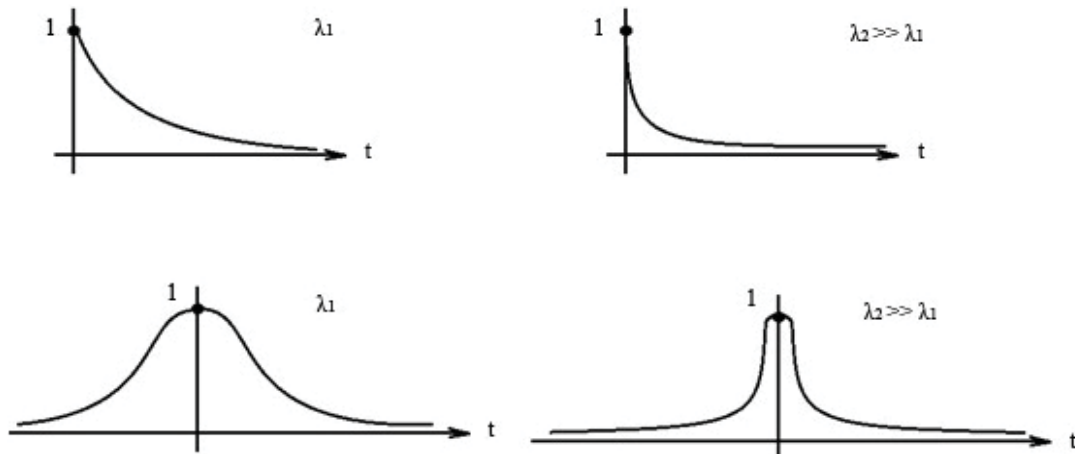


Рис. 1. Графическое отображение ядер

Здесь $K_n(t) \rightarrow 0$ на $]-\infty, +\infty[$. Введем масштаб 1:2, то есть положим $Q_n(t) = 2^{+n} K_n(t)$. Графики $Q_n(t)$ имеют при $n \rightarrow \infty, n \in \mathbb{N}$ характер, изображенный на рис. 2.

Естественно рассмотреть вначале случай одного фиксированного пика фиксированной высоты. Полезно различать две возможности:

а) основание пика (то есть узел асимптотики) находится в крайней точке промежутка интегрирования;

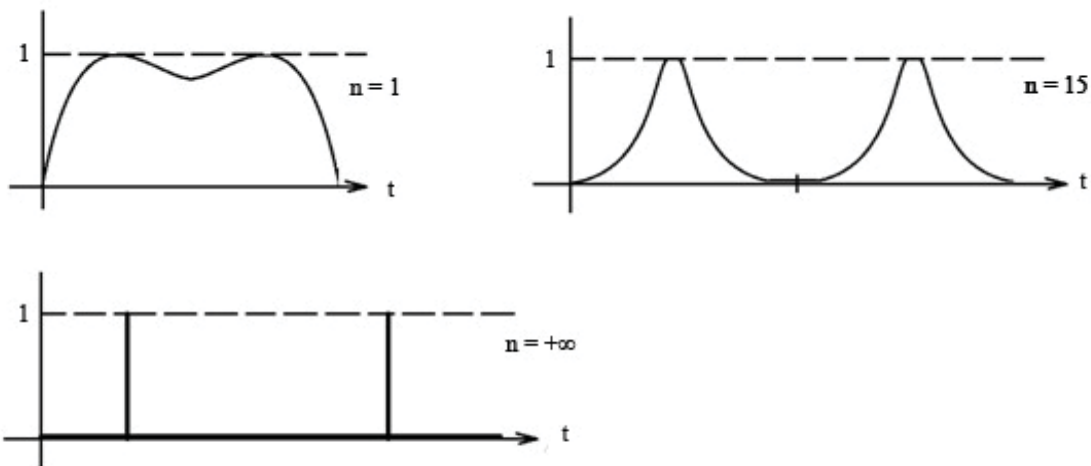


Рис. 2. Графическое отображение специальных ядер

Как уже неоднократно говорилось, содержание метода Лапласа богаче любого количества посвященных этому методу теорем. Можно представить себе ситуацию, когда с ростом λ .

б) узел асимптотики лежит внутри промежутка интегрирования.

Представляет интерес еще одно свойство интеграла Лапласа $\Phi(\lambda) = \int_{\alpha}^{\beta} e^{-\lambda t} f(t) dt$.

Пусть этот интеграл сходится при некотором значении $\lambda = \lambda_1$. Тогда он сходится и при всех λ таких, что $\lambda = \lambda_1$. Нижняя грань λ_0 всех λ , при которых $\Phi(\lambda)$ сходится, называется абсциссой сходимости $\Phi(\lambda)$ (в отличие от степенных рядов, у интегралов $\Phi(\lambda)$ абсциссы сходимости и абсолютной сходимости, в общем случае, различны). В дальнейшем предполагается, что у всех рассматриваемых интегралов Лапласа существует конечная абсцисса сходимости. Следующая вспомогательная лемма позволяет раз и навсегда ограничиться интегралами, распространенными на конечный отрезок.

Лемма. Пусть функция $\Phi(t)$ интегрируема на любом промежутке $]0, \tau[$, $\tau > 0$, интеграл

$$\Phi(\lambda) = \int_0^{+\infty} e^{-\lambda t} \varphi(t) dt$$

сходится при $\lambda = \lambda_0$ и $\delta > 0$ – произвольное положительное число [6].

Тогда для интеграла

$$\Phi_\delta(\lambda) = \int_\delta^{+\infty} e^{-\lambda t} \varphi(t) dt$$

справедлива асимптотическая оценка $\Phi_\delta(\lambda) = \sigma(e^{-\delta})$, $\lambda \rightarrow +\infty$.

Доказательство. По условию интеграл

$$F_0(t) = \int_\delta^t e^{-\lambda_0 \tau} \varphi(\tau) d\tau$$

сходится при $t \rightarrow +\infty$ к некоторому конечному пределу. Также по условию функция $\varphi(\tau)$ (а с ней и $e^{-\lambda_0 \tau} \varphi(\tau)$), локально интегрируема, вследствие чего $F_0(t)$ непрерывная функция при $t \in [0, +\infty[$. Непрерывная на $[0, +\infty[$ функция, имеющая при $t \rightarrow +\infty$ конечный предел, ограничена на $[0, +\infty[$: поэтому существует и конечное число

$$M = \sup_{0 \leq t < +\infty} |F_0(t)|.$$

Заметив это, проведем в интеграле $\Phi_\delta(\lambda)$ при любом $\lambda = \lambda_0$ следующие преобразования:

$$\begin{aligned} \Phi_\delta(\lambda) &= \int_0^{+\infty} e^{-\lambda t} \varphi(t) dt = \\ &= \int_0^{+\infty} e^{-(\lambda - \lambda_0)t} e^{-\lambda_0 t} \varphi(t) dt = \\ &= \left\{ e^{-(\lambda - \lambda_0)t} = u; du = -(\lambda - \lambda_0)e^{-(\lambda - \lambda_0)t} dt \right\} = \\ &= e^{-(\lambda - \lambda_0)t} F_0(t) \Big|_{t=\delta}^{t=+\infty} + \\ &+ (\lambda - \lambda_0) \int_\delta^{+\infty} e^{-(\lambda - \lambda_0)t} F_0(t) dt. \end{aligned}$$

Теперь оценка по модулю дает ($F_0(\delta) = 0$ по определению, $|F_0(+\infty)| \leq M$):

$$|\Phi_\delta(\lambda)| \leq (Me^{\lambda_0 \delta})e^{-\delta \lambda} +$$

$$+ (\lambda - \lambda_0)M \int_\delta^{+\infty} e^{-(\lambda - \lambda_0)t} dt = (2Me^{\lambda_0 \delta})e^{-\delta \lambda},$$

что и требовалось доказать.

Таким образом, любой «хвост» $\Phi_\delta(\lambda)$ интеграла Лапласа с конечной абсциссой сходимости всегда экспоненциально мал. Поэтому с точки зрения асимптотических оценок можно ограничиться любым конечным промежутком $]0, a[$, $a > 0$. Кроме того, эта асимптотика определяется лишь поведением функции $\varphi(t)$ при $t \rightarrow +\infty$.

Предложена математическая модель непараметрической оценки плотности вероятности в условиях малых выборок и метод Лапласа в асимптотической оценке применяемых ядер обеспечивают возможность технического моделирования, без использования реальных технологических процессов и дорогостоящего оборудования. Реализация «обхода» проблем малых выборок обеспечивает увеличение объема исходных данных и повышает эффективность технического моделирования [7].

Список литературы

1. Лапко А.В. Непараметрические системы обработки неоднородной информации / А.В. Лапко, В.А. Лапко. – Новосибирск: Наука, 2007. – 174 с.
2. Parzen E. On the estimation of a probability density function and mode // Ann. Math. Statist. – 1962. – P. 1065.
3. Епаничников В.А. Непараметрическая оценка многомерной плотности вероятности // Теория вероятности и ее применения. – 1969. – Т. 14, Вып. 1.
4. Балан Н.Н., Васин В.А., Ивашов Е.Н., Костомаров П.С., Степанчиков С.В. Моделирование процессов иммерсионной ультрафиолетовой литографии на этапе аванпроектирования // Нелинейный мир. – 2012. – Т. 10, № 7. – С. 454-459.
5. Оливер Ф. Введение в асимптотические методы и специальные функции – М.: Наука, 1978. – 357 с.
6. Садыхов В.Э. Элементы асимптотических методов. – Баку: АЗПИ им. Ч. Ильдрима, 1985. – 141 с.
7. Балан Н.Н., Васин В.А., Ивашов Е.Н., Костомаров П.С., Степанчиков С.В. Оптимизация процессов иммерсионной ультрафиолетовой литографии в условиях малых выборок. В кн.: Фундаментальные проблемы радиоэлектронного приборостроения. Материалы Международной научно-технической конференции «INTERMATIC – 2012», 3–7 декабря 2012 г., Москва / отв. ред.: А.С. Сигов. Ч. 3: Материалы и технологии. – М.: МГТУ МИРЭА – ИРЭ РАН, 2012. – С. 112-116.