

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

И

ПРОЦЕССЫ УПРАВЛЕНИЯ

№ 1, 1998

Электронный журнал,
рег. № П23275 от 07.03.97

<http://www.neva.ru/journal>
e-mail: diff@osipenko.stu.neva.ru

фильтрация и идентификация

АСТАТИЧЕСКИЕ ПОТЕНЦИАЛЬНЫЕ ФУНКЦИИ И ТОЧНОСТЬ ОЦЕНОК МЕТОДА ЛОКАЛЬНОЙ АППРОКСИМАЦИИ

Т.В.Варядченко, А.А.Первозванский, Россия, 195251, Санкт-Петербург,
ул. Политехническая, д.29

С.-Петербургский государственный технический университет

Кафедра "Механика и процессы управления"

e-mail: control1@citadel.stu.neva.ru

Аннотация.

Получены оценки точности оценок Надарая-Ватсона в зависимости от размерности пространства, порядка астатизма функции локальности и параметра класса оцениваемой функции. Также получены формулы для вычисления астатических потенциальных функций в R_n произвольного порядка. Приводятся результаты контрольных расчетов.

1. Введение

Понятие потенциальной функции было введено в [1] для описания процедур обучения в задачах аппроксимации и классификации. В широком смысле потенциальной именуется любая функция, убывающая при удалении от фиксированной точки, источника потенциала. При использовании такой

функции в методе локальной аппроксимации (МЛА) [4] для аппроксимации функции $y(x)$ этой фиксированной точкой является текущий центр аппроксимации x . Дадим здесь определение оценок МЛА и потенциальных функций, следуя [4].

Определение 1 Функция $h(u)$, $u \in R^d$ называется потенциальной функцией с астатизмом порядка α , $h(u) \in S_\alpha$, если

$$\int_{R^d} h(u) du = 1 \quad |h(u)| \leq C_n \exp\left(-\sum_{j=1}^d |u_j|\right) \quad \int_{R^d} h(u) p_s(u) du = 0, \quad s = 1, \dots, \alpha \quad (1)$$

где $p_s(u)$ - произвольный полином степени s от компонент u .

Определение 2 Оценка

$$\hat{y}_n(x, \delta) = \hat{c}^\top(x, \delta) \varphi(0) \quad (2)$$

$$\hat{c} = \arg \min_c \frac{1}{n} \sum_{s=1}^n h\left(\frac{x - x^{(k)}}{\delta}\right) (y^{(k)} - \hat{y}(x^{(k)}, \delta))^2 \quad (3)$$

$$\hat{c}^\top(x, \delta) = \Phi^{-1}(x, \delta) \sum_{k=1}^n h\left(\frac{x - x^{(k)}}{\delta}\right) \varphi(x - x^{(k)}) \varphi^\top(x - x^{(k)}) y^{(k)} \quad (4)$$

$$\Phi(x, \delta) = \sum_{k=1}^n h\left(\frac{x - x^{(k)}}{\delta}\right) \varphi(x - x^{(k)}) \varphi^\top(x - x^{(k)}), \quad (5)$$

где $\varphi(x)$ - заданные функции n переменных, $h(x)$ - потенциальные функции весов, зависящие от некоторого управляющего параметра δ , называется оценкой МЛА.

В случае нулевой степени аппроксимации, когда $\varphi(x) = 1$, оценки (2 - 5) записываются в виде

$$\hat{y}_n(x, \delta) = \frac{\sum_{i=1}^n y(x^{(k)}) h\left(\frac{x - x^{(k)}}{\delta}\right)}{\sum_{i=1}^n h\left(\frac{x - x^{(k)}}{\delta}\right)} \quad (6)$$

и называются оценками Надарая-Ватсона. Для этого случая оценки вычисляются наиболее просто.

Исследование точности оценок (2) - (5) проведено неполно [2-5]. Следует отметить утверждение о вероятностной сходимости аппроксимации (2 - 5) при неограниченном возрастании числа элементов обучающей выборки и стремлении к нулю параметра локальности [1,3], а также оценки точности сверху, данные в [4] при фиксированных объемах обучающей выборки.

Что касается возможностей конструктивного построения потенциальных функций, то только некоторые формулы, взятые из разных источников для различных частных случаев приведены там же [4]. В данной работе получены формулы, позволяющие построить потенциальные функции произвольного порядка астатизма в R^d конкретного вида и даны конструктивные оценки точности оценок (6) в зависимости от порядка астатизма, размерности пространства d , а также от параметра p класса оцениваемой функции. Построить такие же конструктивные оценки точности в общем случае оценок МЛА пока представляется затруднительным и будет являться предметом дальнейших исследований.

Введем также в рассмотрение определение класса оцениваемой функции в виде [7].

Определение 3.

Пусть $f(x), x = (x_j) \in X \subset R^d$ такова, что

$$|D_t^s f(x)| \leq L_f, s = 0, \dots, p, \quad |D_t^p f(x) - D_t^p f(z)| \leq L_f \sum_{j=1}^d |x_j - z_j|, x, z \in X, \quad (7)$$

где

$$D_t^s f = \frac{\partial^s f}{\prod_j (\partial x_j) t_j}, \quad \sum_{j=1}^d t_j = s, \quad t = (t_j) \quad (8)$$

Тогда

$$f(x) \in H(p), \quad \|f(x)\|_{H(p)} = L_f \quad (9)$$

2. Основные результаты

Основные результаты работы заключаются в следующих утверждениях.

Теорема Пусть $x^{(k)} \in C^d$, $C^d = \{x/0 \leq x_j \leq 1, j = 1, \dots, d\}$ и $w_k = \delta^{-d} y(x^{(k)})$, тогда при достаточно малых $\delta > 0$ для любой оценки потенци-

ального типа $\hat{y}_n(x, \delta) = \sum \delta^{-d} y(x^{(k)}) h\left(\frac{x-x^{(k)}}{\delta}\right)$

$$I_n(\delta) = \sup_{x \in C_1^d} |y(x) - \hat{y}_n(x, \delta)| \leq A_0 \delta^{r_0} + A_1 \delta^{-r_1} g(n), \quad (10)$$

где

$$C_1^d = \{x \quad / \quad \delta^{1-\epsilon} \leq x_j \leq 1 - \delta^{1-\epsilon}\}, \quad 1 \geq \epsilon \geq 0, \quad (11)$$

величины A_0, A_1 не зависят от δ и n , целые положительные показатели r_0, r_1 зависят только от класса Функций $y(x)$ и порядка астатизма $h(u)$, так что

$$r_0 = \bar{p} + 1, \quad \bar{p} = \min(p, \alpha), \quad r_1 = p$$

если $y(x) \in H(p), h(u) \in S_\alpha$, а $h(u)$ удовлетворяет только условию нормировки (1).

Функции $g(n)$ зависят от гладкости $y(x)$ и степени неравномерности сетки $\{x^{(k)}\}$ на C^d .

Могут быть построены сетки, для которых

$$g(n) = O(n^{-\frac{p+1}{d}}), \quad \text{если } y(x) \in H(p),$$

Следствие При фиксированной сетке может быть выбрано такое $\delta = \delta^*$, что

$$I_n(\delta^*) \leq A[g(n)]^{\frac{r_0}{r_0+r_1}}, \quad \delta^* = O\left(g(n)^{\frac{1}{r_0+r_1}}\right),$$

а при наилучших сетках

$$I_n(\delta^*) = O\left(n^{-\frac{r_0(p+1)}{(r_0+r_1)d}}\right), \quad \text{если } y(x) \in H(p), \quad (12)$$

Доказательство приводится полностью в [6]. Для оценки Надарая-Ватсона справедливы все утверждения теоремы 1.

Доказательство приведено в [6] и следует из того, что знаменатель в (6) (с точностью до множителя n^{-1}) является квадратурой интеграла

$$\begin{aligned} \int_{C^d} h\left(\frac{x-u}{\delta}\right) du &= \int_{R^d} h\left(\frac{x-u}{\delta}\right) du + \Delta_h(\delta, x) = \\ &= \delta^d \int_{R^d} h(u) du + \Delta_h(\delta, x) = \delta^d [1 + \delta^{-d} \Delta_h(\delta, x)]. \end{aligned}$$

где поправка Δ_h соответствующим образом оценивается.

Теорема 2 Астатические функции произвольного порядка $(2\alpha - 1)$ в R^d определяется соотношениями

$$h(u) = \frac{1}{(2\pi)^{d/2}} \sum_{i=1}^{\alpha} c_i \|u\|^{2(i-1)} \exp(-\|u\|^2/2) \quad (13)$$

где вектор коэффициентов $c^T = (c_1, \dots, c_\alpha)$ является решением системы уравнений

$$Ac = e, \quad a_{i1} = 1, i = 1, \dots, \alpha$$

,

$$a_{kl} = \prod_{j=k-2}^{l-2} (d + 2j), k = 2, \dots, \alpha, l = 2, \dots, \alpha, \quad e^T = (1, 0, \dots, 0) \quad (14)$$

В частности, запишем вид матрицы A для функций 7-го порядка астатизма

$$A = \begin{pmatrix} 1 & d & d(d+2) & d(d+2)(d+4) \\ 1 & d+2 & (d+2)(d+4) & (d+2)(d+4)(d+6) \\ 1 & d+4 & (d+4)(d+6) & (d+4)(d+6)(d+8) \\ 1 & d+6 & (d+6)(d+8) & (d+6)(d+8)(d+10) \end{pmatrix} \quad (15)$$

Формула (15) дает полное представление о структуре матрицы для оператора произвольного порядка астатизма.

Доказательство следует из непосредственной записи соотношений астатизма (1) для функций указанного вида. При этом используются значения интегралов [9]

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx &= 1 \\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2k} \exp\left(-\frac{x^2}{2}\right) dx &= 1 \cdot 3 \cdot 5 \cdots (2k - 1), k = 1, 2, \dots \\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2k+1} \exp\left(-\frac{x^2}{2}\right) dx &= 0, \quad k = 0, \dots \end{aligned}$$

Также следует иметь в виду, что соотношения

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} u_i^{2k} h(u) du = 0$$

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} u_i^{2l} u_j^{2m} h(u) du = 0$$

для всех $i, j = 1, \dots, d$ дают одно уравнение для каждого $k = l + m, k = 1, \dots, (\alpha - 1)$ в систему уравнений для определения коэффициентов c . Иными словами соответствующее уравнение следует при каждом $k = 1, \dots, (\alpha - 1)$ из соотношения

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} u_1^{2k} h(u) du = 0$$

Замечание Приведенные функции (13) являются функциями расстояния $r = \| u \|^2$.

3. Результаты контрольных расчетов

Ниже приводятся результаты расчетов, полученные на сетках Хаммерсли [8] на отрезке или в квадрате $(0,1)$ при аппроксимации функций одной $d = 1$ и двух $d = 2$ переменных с весовыми функциями вида (13)-(14), зависящими от расстояния $r = \| u \|^2$ и имеющими первый и третий порядок астатизма

$$h^{(1)}(u) = \frac{1}{\sqrt{2\pi}} \exp(- \| u \|^2 / 2) \tag{16}$$

$$h^{(3)}(u) = \frac{1}{(2\pi)^{-3/2}} (3/2 - 1/2 \| u \|^2) \exp(- \| u \|^2 / 2) \tag{17}$$

на тестовых функциях вида "параболы" и "сужающийся фарватер"

$$f(x, y) = (x - 0.5)^2 + 0.5(y - 0.5)^2 \tag{18}$$

$$f(x, y) = \exp(-y(x - 0.5)^2), \quad \| u \|^2 = x^2 + y^2 \tag{19}$$

В одномерном случае в качестве тестовой функции взята парабола $f(x) = x^2, \| u \|^2 = x^2$. Результаты для одномерного случая приведены в таблицах 1 и 2. В таблице 1 — для более низкого порядка астатизма весовой функции (16), в таблице 2 — для более высокого (17).

Таблицы 3,4 содержат результаты для параболы (18), таблица 3 — для более низкого порядка астатизма весовой функции (16), таблица 4 — для более высокого (17).

В таблицах 5,6 приведены результаты, полученные на модельной функции вида (19) также соответственно для весовых функций более низкого и более высокого порядка астатизма указанного вида.

Обозначения в таблицах соответствуют обозначениям теоремы 1. Как видно, результаты, приведенные в таблицах 1-6, подтверждают основное утверждение теоремы 1.

Таблица 1.

N	δ^*	$I^*(N)$
10	0.19	0.034770
50	0.03	0.015419
250	0.01	0.005376
500	0.002	0.002342
750	0.001	0.000994

Таблица 2.

N	δ^*	$I^*(N)$
10	0.09	0.023779
50	0.04	0.005961
250	0.02	0.001689
500	0.01	0.000869
750	0.002	0.000367

Таблица 3.

N	δ^*	$I^*(N)$
100	0.06	0.0537
900	0.04	0.0348
2500	0.03	0.0176
6400	0.01	0.0076

Таблица 4.

N	δ^*	$I^*(N)$
100	0.08	0.0289
900	0.05	0.0174
2500	0.04	0.0074
6400	0.04	0.0048

Таблица 5.

N	δ^*	$I^*(N)$
100	0.07	0.0392
900	0.03	0.0278
2500	0.03	0.0140
6400	0.01	0.0058

Таблица 6.

N	δ^*	$I^*(N)$
100	0.09	0.0193
900	0.05	0.0147
2500	0.04	0.0049
6400	0.05	0.0035

На рис.1 приведены трехмерные образы функций оценок, полученные смещением сечений оценок вдоль оси y для функции (19) при оптимальном значении параметра $\delta = \delta^*$ и значении параметра $\delta < \delta^*$, при котором оценка перестает быть гладкой.

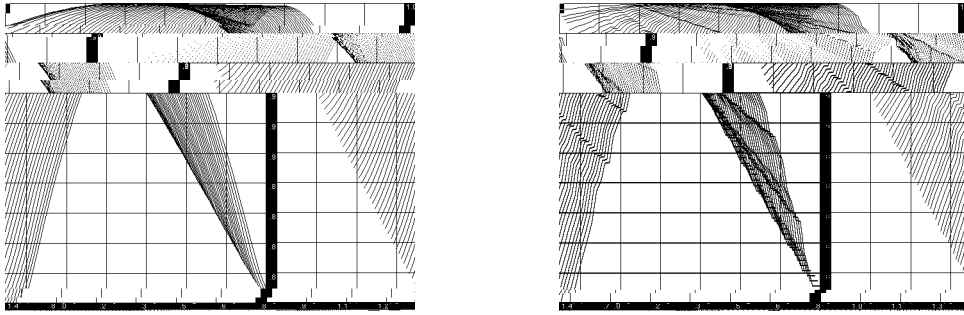


Рис.1. Оценка при оптимальном значении параметра (слева) и значении, меньшем оптимального (справа).

4. Заключение

Полученные формальные результаты и их подтверждение в результате контрольных расчетов позволяют дать конкретные рекомендации по построению оценок локального типа при аппроксимации функций.

Основные практические выводы заключаются в следующем:

- если задаваемый допуск на точность равен Δ , то число элементов обучающей выборки $\{x^k, y^k\}$ должно определяться соотношением $n = \Delta^q$, где q зависит от параметра гладкости p аппроксимирующей функции, порядка астатизма α потенциальной функции и размерности пространства d . Так, если выбрать $\alpha = p$, то

$$q = \frac{d(2\alpha + 1)}{(\alpha + 1)^2}$$

Ясно, что при повышении требований к точности объем выборки измерений быстро растет с размерностью пространства, если этот рост не компенсировать увеличением степени астатизма используемой потенциальной функции. В частности, для классической потенциальной функции типа гауссовской кривой

$$h(u) = 1/(\sqrt{2\pi})^d \exp\left(-\frac{u^2}{2}\right)$$

степень астатизма $\alpha = 1$, и такая функция неэффективна при аппроксимации функций большого числа переменных. Для гладких функций жела-

тельно использовать при аппроксимации потенциальные функции с высокой степенью астатизма. Результаты теоремы 2 позволяют построить такие функции конкретного вида;

- при конечном объеме измерений, вообще говоря, не гарантируется достижение предельной точности аппроксимации. Если последовательность измерений $\{x^k, y^k\}$ не представительна, т.е. из нее невозможно выбрать элементы, образующие достаточно "хорошую" сетку, то точность будет заведомо плохой. "Хорошие" сетки, рекомендуемые теорией квадратур, состоят из точек, равномерно расставленных на области аппроксимации (в частности, гиперкубе). Таковы использованные в контрольных расчетах сетки Хаммерсли. При этом равномерность отнюдь не означает "регулярную" расстановку узлов сетки с одинаковым шагом по всем переменным, а близка к псевдослучайным равномерно распределенным последовательностям. Таким образом, предельные возможности по достижимой точности сжимаются с ростом размерности пространства и недостаточной представительности обучающей последовательности по области аппроксимации;

- если все параметры α, p и d заданы и выборка представительна в описанном смысле, допуск по точности Δ может быть просто вычислен по формуле (12).

Список литературы

- [1] Айзерман М.А., Броверман Э.М., Розоноер Л.В. Метод потенциальных функций в теории обучения машин. М.: Наука. 1970.
- [2] Barron A.R. Universal approximation bounds for superposition of a sigmoidal function // IEEE Tr. Inf.Th., IT-39, 1993, p.930–945.
- [3] Girosi F., Poggio T. Networks and the best approximation property // Biological Cyb., v. 63, 1990, p.169–176.
- [4] Катковник В.Я. Непараметрическая идентификация и сглаживание данных: метод локальной аппроксимации. М.: Наука, 1985.
- [5] Первозванский А.А., Буцев А.В. Локальная аппроксимация с помощью ИНС // Автоматика и телемеханика, 1995.
- [6] Первозванский А.А. Оценки точности аппроксимации с помощью суперпозиции потенциальных функций // Проблемы передачи информации. N4. 1996.

- [7] Бахвалов Н.Н. О приближенном вычислении многомерных интегралов // Вестник МГУ, мат. сер. N 4,1959, с.3–18.
- [8] Соболев И.М. Многомерные квадратурные формулы и функции Хаара // М.: Наука, 1969.
- [9] Бронштейн И.Н., Семендяев К.А. Справочник по математике для инженеров и учащихся втузов. М.:Наука, 1986.