Записки научных
семинаров ПОМИ
Том 412, 2013 г.

И. А. Ибрагимов

ОБ ОДНОМ ОБОБЩЕНИИ ПРОЕКЦИОННЫХ ОЦЕНОК Н. Н. ЧЕНЦОВА

1. Проекционные оценки впервые появились в 1962 г. в работе Н. Н. Ченцова [1], как метод построения оценок бесконечномерного, например, функционального параметра. В общих чертах этот метод заключается в следующем. Допустим, что подлежащий оценке параметр $\theta \in \Theta \subseteq H$, H – гильбертово пространство с нормой $\|\cdot\|$, и качество оценки $\widehat{\theta}$ определяется величиной $\mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2$. Метод Ченцова рекомендует рассмотреть конечномерное подпространство H_N пространства H, хорошо приближающее Θ, и оценивать не сам параметр θ, а его проекции θ_N на H_N. Последняя задача уже есть задача оценки конечномерного параметра.

Для определенности рассмотрим следующую задачу, в которой метод построения проекционных оценок выглядит особенно просто и естественно. Функция $f \in F \subset L_2(0,1)$ наблюдается в гауссовском белом шуме интенсивности ε. Это означает, что в качестве наблюдения мы имеем

$$X_\varepsilon(t) = \int_0^t f(s)\,ds + \varepsilon w(t), \tag{1.1}$$

где $w(t)$ – процесс Винера.

Пусть H_N – конечномерное подпространство $L_2([0,1])$ размерности N, а $\varphi_1,\dots,\varphi_N$ – ортогональный базис в H_N. Проекция f на H_N имеет вид

$$f_N = \sum_1^N a_j\varphi_j, \quad a_j = (f,\varphi_j) = \int_0^1 f(t)\varphi_j(t)\,dt,$$

Ключевые слова: проекционные оценки Ченцова, непараметрическая теория оценивания, ядерные оценки, воспроизводящие ядра.

Работа выполнена при частичной поддержке РФФИ, грант 11-01-00577-а, “Ведущие научные школы”, грант НШ-1216.2012.1, и Программы фундаментальных исследований РАН “Современные проблемы теоретической математики”.

и естественной оценкой для f_N является статистика

$$\widehat{f}_N = \sum_1^N \widehat{a}_j \varphi_j, \quad \widehat{a}_j = \int_0^1 \varphi_j(t)\, dX_\varepsilon(t).$$

Очевидно, что

$$\mathbf{E}_f \|f - \widehat{f}_N\|^2 = \sum_1^N \mathbf{E}_f |a_j - \widehat{a}_j|^2 + \|f - f_N\|^2 \leqslant N\varepsilon^2 + \big(\operatorname{dist}(F, H_N)\big)^2.$$

Поэтому для любого $\delta > 0$ можно найти такую оценку $\widehat{f}_\delta$, что

$$\sup_{f \in F} \mathbf{E}_f \|f - \widehat{f}_\delta\| \leqslant N\varepsilon^2 + (1+\delta) d_N^2(F),$$

$$\inf_{\widehat{f}} \sup_{f \in F} \mathbf{E}_f \|f - \widehat{f}\|^2 \leqslant N\varepsilon^2 + d_N^2(F). \tag{1.2}$$

Здесь $d_N(F)$ означает N-й поперечник Колмогорова множества F в $L_2([0,1])$. Поперечники множеств нормированных пространств были введены А. Н. Колмогоровым [2]. Относительно поперечников см., например, [3–5].

2. В работе [6], хотя и не совсем явно, было отмечено, что в определенных случаях проекции на бесконечномерные пространства с ограниченным воспроизводящим ядром допускают столь же хорошие оценки, что и проекции на конечномерные пространства. Этот феномен и основанные на нем методы оценивания исследовался в работе [7]. Настоящая работа продолжает это исследование.

Напомним определение пространств с воспроизводящим ядром (см. [8, с. 139], [9]). Пусть $\mathfrak{X}$ – некоторое абстрактное множество точек x и пусть система H комплексных функций $f(x)$, заданных на $\mathfrak{X}$, образует гильбертово пространство со скалярным произведением (f, g). Комплексная функция $K(x, y)$, определенная на $\mathfrak{X} \times \mathfrak{X}$, называется воспроизводящим ядром для H, если

a) для любого фиксированного $y \in \mathfrak{X}$ функция $K(x, y)$ как функция от x есть элемент H;

b) для любого элемента $f \in H$

$$f(y) = (f(\cdot), K(\cdot, y))$$

и, следовательно,

$$\overline{f(y)} = (K(\cdot, y), f(\cdot)).$$

В силу условия b),

$$K(x,y) = (K(\cdot,y), K(\cdot,x)).$$

Если H имеет воспроизводящее ядро, то это ядро единственно.

Воспроизводящее ядро положительно определено. В частности,

$$K(x,y) = \overline{K(y,x)}, K(x,x) \geqslant 0, \quad |K(x,y)|^2 \leqslant K(x,x)K(y,y).$$

Если $\{\varphi_j(x)\}$ – это какая-нибудь ортонормальная система, то

$$\sum |\varphi_j(x)|^2 \leqslant K(x,x).$$

Если же пространство H сепарабельно, то оно имеет воспроизводящее ядро в том и только том случае, если для любого ортонормированного базиса $\{\varphi_j\}$ сумма $\sum |\varphi_j(x)|^2 < \infty$. При этом

$$K(x,y) = \sum \varphi_j(x)\overline{\varphi_j(y)}. \tag{2.1}$$

Отметим еще, что пространство H имеет воспроизводящее ядро в том и только том случае, если все линейные функционалы $L_x(f) = f(x)$ непрерывны. При этом

$$|f(x)| \leqslant \big(K(x,x)\big)^{1/2} \cdot \|f\|.$$

Ниже мы рассматриваем только вещественные H, так что $K(x,y) = K(y,x)$.

Примеры.

2.1. Всякое конечномерное пространство H есть пространство с вопроизводящим ядром $K(x,y) = \sum\limits_1^N \varphi_j(x)\varphi_j(y)$, где $\{\varphi_j\}$ – ортонормальный базис в H.

2.2. Пусть G – измеримое подмножество R^d, имеющее конечную положительную лебегову меру. Обозначим $H_G = H$ подпространство $L^2(R^d)$, состоящее из функций f, преобразование Фурье которых

$$\widetilde{f}(t) = \int\limits_{R^d} e^{i(t,x)} f(x)\,dx$$

обращается в нуль вне множества G. Тогда H_G – пространство с воспроизводящим ядром

$$K(x,y) = \frac{1}{(2\pi)^d} \int\limits_G e^{i(t,x-y)}\,dt. \tag{2.2}$$

Действительно, если $f \in H_G$, то

$$\int\limits_{R^d} K(x,y) f(x)\,dx = \frac{1}{(2\pi)^d} \int\limits_G e^{-i(t,y)}\,dt \int\limits_{R^d} e^{i(t,x)} f(x)\,dx$$
$$= \frac{1}{(2\pi)^d} \int\limits_G e^{-i(t,y)} \widetilde{f}(t)\,dt = f(y).$$

2.3. Пусть интервалы $[a_j, a_{j+1})$ не пересекаются. Пусть функции $\varphi_j \in L_2(R^1)$ обращаются в нуль вне интервала (a_j, a_{j+1}) и $\|\varphi_j\| = 1$. Пусть H – подпространство $L_2(R^1)$, натянутое на векторы $\{\varphi_j\}$. Тогда H имеет воспроизводящее ядро

$$K(x,y) = \sum_j \varphi_j(x)\varphi_j(y). \tag{2.3}$$

3. Ниже мы будем рассматривать задачу оценивания параметра $\theta \in \Theta \subseteq L_2(\mathfrak{X}, \mathfrak{A}, \mu) = L_2$, μ – σ-конечная мера на σ-алгебре $\mathfrak{A}$ подмножеств $\mathfrak{X}$.

Мы рассмотрим три следующие статистические задачи.

I. Оценка плотности распределения. Наблюдениями в этой задаче являются независимые одинаково распределенные $\mathfrak{X}$-значные случайные величины

$$X_1, \ldots, X_n$$

с неизвестной плотностью распределения $\theta(x)$ относительно μ. Неизвестная плотность θ принадлежит известному множеству $\Theta \subseteq L_2$. Качество оценки $\widehat{\theta}$ измеряется величиной $\mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2$.

II. Оценка функции в аддитивном гауссовском шуме. Для любого $\varphi \in L_2$ наблюдается гауссовская случайная величина

$$X(\varphi) = (\theta, \varphi) + \xi(\varphi) = \int\limits_{\mathfrak{X}} \theta(x)\varphi(x)\,d\mu + \xi(\varphi),$$

где функция $\theta \in \Theta$ подлежит оценке. Величины $\xi(\varphi)$ – гауссовские с

$$\mathbf{E}\xi(\varphi) = 0, \quad \mathbf{E}\xi(\varphi)\xi(\psi) = \int\limits_{\mathfrak{X}} \varphi(x)\psi(x)\sigma^2(x)\,d\mu.$$

III. Оценка плотности интенсивности пуассоновского множества Π. На множестве $\mathfrak{X}$ наблюдается случайное пуассоновское множество Π с неизвестной плотностью интенсивности θ относительно μ. Плотность $\theta \in \Theta \subseteq L_2$. Надлежит по наблюдениям Π оценить θ.

Следуя указанному выше методу, мы будем сперва оценивать проекции θ на пространство H_K с воспроизводящим ядром K, $H_K \subseteq L_2$.

4. Мы начнем с исследования более простой задачи II.

Теорема 4.1. *Пусть $H_K \subseteq L_2$ – подпространство с воспроизводящим ядром $K(x,y)$. Пусть $\Theta \subseteq H_K$. Оценим θ статистикой*

$$\widehat{\theta}_K(y) = X(K(\cdot, y)).$$

Тогда

$$\mathbf{E}_\theta \|\widehat{\theta}_K - \theta\|^2 \leqslant \int\limits_{\mathfrak{X}} K(x,x)\sigma^2(x)\, d\mu \tag{4.1}$$

и, следовательно,

$$\inf_{\widehat{\theta}} \sup_{\theta \in \Theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 \leqslant \int\limits_{\mathfrak{X}} K(x,x)\sigma^2(x)\, dx. \tag{4.2}$$

Доказательство. Так как $\theta \in H_K$, то

$$\theta(y) = \big(\theta, K(\cdot, y)\big) = \int\limits_{\mathfrak{X}} \theta(x) K(x,y)\mu(dx).$$

Поэтому $X(K(\cdot,y)) = \theta(y) + \xi(K(\cdot,y))$ и, следовательно,

$$\begin{aligned}
\mathbf{E}_\theta \|\widehat{\theta}_K - \theta\|^2 &= \mathbf{E} \int\limits_{\mathfrak{X}} |\xi(K(\cdot,y)|^2 \mu(dy) \\
&= \int\limits_{\mathfrak{X}} \mu(dy) \int\limits_{\mathfrak{X}} K(x,y)K(x,y)\sigma^2(x)\mu(dx) \\
&= \int\limits_{CX} \sigma^2(x)\mu(dx) \int\limits_{\mathfrak{X}} K(x,y)K(x,y)\mu(dy) \\
&= \int\limits_{\mathfrak{X}} \sigma^2(x) K(x,y)\mu(dx).
\end{aligned}$$

□

Следующая теорема показывает, что неравенство (4.2) теоремы 4.1, вообще говоря, нельзя улучшить.

Теорема 4.2. *Пусть в условиях теоремы* 4.1 *множество* $\Theta = H_K$. *Тогда*

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 = \int\limits_{\mathfrak{X}} K(x,x)\sigma^2(x)\,dx. \tag{4.3}$$

Доказательство. Пусть $\{\varphi_j\}$ – ортонормальный базис в H_K. Обозначим L_N подпространство H_K, натянутое на векторы $\varphi_1,\dots,\varphi_N$. Очевидно,

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 \geqslant \inf_{\widehat{\theta}} \sup_{\theta\in L_N} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2. \tag{4.4}$$

Ясно, что при оценивании $\theta \in L_N$ достаточно ограничиться оценками $\widehat{\theta} \in L_N$. Таким образом, полагая $\theta = \sum\limits_1^N a_j\varphi_j$, $\widehat{\theta} = \sum\limits_1^N \widehat{a}_j\varphi_j$, мы найдем, что

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 \geqslant \inf_{\{\widehat{a}_j\}} \sup_{\{a_j\}} \mathbf{E}_a \biggl(\sum_1^N |a_j - \widehat{a}_j|^2\biggr). \tag{4.5}$$

Рассмотрим вспомогательную задачу оценивания N-мерного параметра $a = (a_1,\dots,a_N)$ по наблюдению

$$X = a + \xi,$$

где $\xi = (\xi_1,\dots,\xi_N)$ – гауссовский вектор с $\mathbf{E}\xi = 0$ и корреляционной матрицей

$$R = \|r_{kl}\|, \quad r_{kl} = \mathbf{E}\xi_k\xi_l = \int\limits_{\mathfrak{X}} \varphi_k(x)\varphi_l(x)\sigma^2(x)\,d\mu.$$

Риск оценки $\widehat{a}$ определяется квадратичной функцией потерь $W(z) = \sum\limits_1^N z_i^2$.

Для этой задачи

$$\inf_{\widehat{a}_j} \sup_a \mathbf{E}_a \biggl(\sum_1^N |\widehat{a}_j - a_j|^2\biggr) = \mathbf{E}W(\xi)$$

$$= \operatorname{tr} R = \sum_1^N \int\limits_{\mathfrak{X}} \varphi_j^2(x)\sigma^2(x)\,d\mu. \tag{4.6}$$

Последнее равенство, по-видимому, хорошо известно специалистам. Его можно легко доказать с помощью аргументов, используемых при доказательстве теоремы Гаека (см. [10, теорема 2.12.1]; заметим, что здесь мы имеем дело с "абсолютной", а не локальной и асимптотической нормальностью).

Возвращаясь к основной задаче, мы находим из (4.4)–(4.6), что для всех N

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \geqslant \int\limits_{\mathfrak{X}} \left(\sum_1^N \varphi_j^2(x)\right)\sigma_j^2(x)\,d\mu.$$

Отсюда следует, что

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \geqslant \int\limits_{\mathfrak{X}} \sum_1^\infty \varphi_j^2(x)\sigma^2(x)d\mu = \int\limits_{\mathfrak{X}} K(x,x)\sigma^2(x)\,d\mu.$$

Обратное неравенство есть следствие теоремы 4.1.

Теорема доказана. □

Примеры

4.1. В условиях примера 2.1, если $\Theta = H_K \subseteq L_2$, а $\sigma(x) \equiv \varepsilon$

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 = \varepsilon^2 \sum_1^N \int\limits_{\mathfrak{X}} \varphi_j^2(x)\,d\mu = N\varepsilon^2. \tag{4.7}$$

4.2. В условиях примера 2.2 $K(x,x) = (2\pi)^{-d}\mathrm{mes}\,\mathrm{G}$, так что (ср. ниже теоремы 5.2, 6.2)

$$\inf_{\widehat{\theta}} \sup_{\theta\in N_K} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 = \frac{\mathrm{mes}\,\mathrm{G}}{(2\pi)^d} \int\limits_{R^d} \sigma^2(x)\,dx. \tag{4.8}$$

4.3. Пусть в условиях примера 2.3 $a_j = \frac{j}{N}$, а $\varphi_j(x) = \sqrt{N}$, $\frac{j}{N} \leqslant x \leqslant \frac{j+1}{N}$. Тогда

$$\inf_{\widehat{\theta}} \sup_{\theta\in H_K} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 = N \int\limits_{-\infty}^{\infty} \sigma^2(x)\,dx. \tag{4.9}$$

5. Оценка плотности распределения. В этом пункте мы рассмотрим задачу I в предположении, что $\Theta \subseteq H_K \subseteq L_2(\mathfrak{X}, \mathfrak{A}, \mu)$ для какого-нибудь подпространства H_K.

Теорема 5.1. *Пусть $H_K \subseteq L_2$ – пространство с воспроизводящим ядром $K(x,y)$. Пусть $\Theta \subseteq H_K$. Оценим θ статистикой*

$$\widehat{\theta}_K(y) = \frac{1}{n}\sum_1^n K(X_j, y).$$

Тогда

$$\mathbf{E}_\theta \|\widehat{\theta}_K - \theta\|^2 \leqslant \frac{1}{n}\Big(\int\limits_{\mathfrak{X}} K(x,x)\theta(x)\,d\mu - \|\theta\|^2\Big) \tag{5.1}$$

и, следовательно,

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 \leqslant \frac{1}{n}\sup_\theta \int\limits_{\mathfrak{X}} K(x,x)\theta(x)\,d\mu. \tag{5.2}$$

Доказательство. В силу воспроизводящих свойств ядра K

$$\mathbf{E}_\theta \widehat{\theta}_K(y) = \int\limits_{\mathfrak{X}} K(x,y)\theta(x)\mu(dx) = \theta(y),$$

так что $\widehat{\theta}_K$ – несмещенная оценка θ. Далее,

$$\begin{aligned}
\mathbf{E}_\theta \|\widehat{\theta}_K - \theta\|^2 &= \int\limits_{\mathfrak{X}} \frac{1}{n}\mathbf{E}_\theta\big(K(X_1,y) - \mathbf{E}_\theta K(X_1,y)\big)^2 \mu(dy) \\
&= \frac{1}{n}\int\limits_{\mathfrak{X}} \theta(x)\mu(dx) \int\limits_{\mathfrak{X}} K(x,y)K(x,y)\mu(dy) - \frac{\|\theta\|^2}{n} \\
&= \frac{1}{n}\Big(\int\limits_{\mathfrak{X}} K(x,x)\theta(x)d\mu - \|\theta\|^2\Big).
\end{aligned}$$

Теорема доказана. □

Заметим, что в отличие от предыдущего раздела неравенство (5.2) не является точным даже в случае $\Theta = H_K$, так как в пространстве H_K может не оказаться достаточного числа положительных функций. Пусть, например, $\mathfrak{X} = R^1$ и θ суть плотности распределения по отношению к мере Лебега. Пусть H есть подпространство $L_2(R^1)$, натянутое

на векторы $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $\varphi_j(x) = \frac{\sin 2^j x}{1+x^2}$, $j = 0, \dots, N$. В H лежит ровно одна плотность распределения, так что $\inf\limits_{\widehat{\theta}} \sup\limits_{\theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\| = 0$.

Приведем два примера, когда неравенство (5.2) оказывается точным. В обоих этих примерах $L_2(\mathfrak{X}, \mathfrak{A}, \mu) = L_2(R^d)$, что не оговаривается особо.

Теорема 5.2. *Рассмотрим задачу* I *оценки плотности распределения θ в предположении, что $\theta \in \Theta \subseteq H_K$, H_K есть пространство с воспроизводящим ядром примера* 2.2 *и Θ содержит все плотности распределения из H_K. Тогда*

$$\liminf_n \inf_{\widehat{\theta}} \sup_{\theta} n\mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 = \sup_{\theta} \int\limits_{R^d} K(x,x)\theta(x)\,dx = \frac{\operatorname{mes} \mathrm{G}}{(2\pi)^d}. \tag{5.3}$$

Доказательство. Эта теорема доказана в [6]. □

Теорема 5.3. *Рассмотрим задачу* I *оценки плотности распределения θ в предположении, что $\theta \in \Theta$, $\Theta \subseteq H_K$ и содержит все плотности распределения из H_K, а H_K есть пространство с воспроизводящим ядром примера* 2.3, *причем $a_j = \frac{j}{N}$, $j = \cdots - 1, 0, 1, \dots$, а $\varphi_j(x) = \sqrt{N}$ для $x \in [a_j, a_{j+1})$. Тогда*

$$\liminf_n \inf_{\widehat{\theta}} \sup_{\theta \in \Theta} n\mathbf{E}_\theta \|\theta - \widehat{\theta}\|^2 = \sup_{\theta \in \Theta} \int\limits_{-\infty}^{\infty} K(x,x)\theta(x)\,dx = N. \tag{5.4}$$

Доказательство. Пусть M – фиксированное целое положительное число. Положим $q(x) = \frac{\sqrt{N}}{M+1} \sum\limits_0^M \varphi_j(x)$ и рассмотрим семейство функций

$$\theta(x) = \theta(x; a) = q(x) + \sum_1^M a_j(\varphi_j(x) - \varphi_0(x)), \tag{5.5}$$

зависящих от параметра $a = (a_1, \dots, a_M)$. Очевидно, что

$$\int\limits_{-\infty}^{\infty} \theta(x; a)\,dx = 1.$$

Если $a \in A = \left\{a : |a_j| < \frac{\sqrt{N}}{M+1}\right\}$, то $\theta(x; a) \geqslant 0$. Таким образом, если $a \in A$, то $\theta(x; a)$ есть плотность распределения в R^1. Будем обозначать

это семейство плотностей тоже буквой A. Очевидно, $A \subseteq \Theta$ и потому для любой оценки $\widehat{\theta}$

$$\sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \geqslant \sup_{\theta\in A} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2. \tag{5.6}$$

При оценивании плотности $\theta \in A$ достаточно ограничиться оценками вида $\widehat{b}\cdot q + \sum\limits_1^N \widehat{a}_j(\theta_j-\theta_0)$, где $\widehat{b}$, $\widehat{a}_j$ – какие-то статистики выборки $X_1,\dots,X_n$. Для таких оценок θ

$$\|\widehat{\theta}-\theta\|^2 = |\widehat{b}-1|^2\,\|q\|^2 + \left\|\sum_1^N (\widehat{a}_j-a_j)(\theta_j-\theta_0)\right\|^2,$$

так что можно положить $\widehat{b}=1$ и рассматривать лишь оценки вида $q+\sum\limits_1^N \widehat{a}_j(\theta_j-\theta_0)$. Таким образом

$$\sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \geqslant \inf_{\widehat{\theta}} \sup_{\theta\in A} \mathbf{E}_\theta \left\{ \sum_1^N |\widehat{a}_j-a_j|^2 + \left|\sum_1^N (\widehat{a}_j-a_j)\right|^2 \right\}.$$

Чтобы оценить правую часть последнего неравенства, рассмотрим параметрическую задачу оценивания параметра a плотности распределения $\theta(x;a)$ по наблюдениям $X_1,\dots,X_n$. Обозначим через $I_M(a) = I(a) = \|I_{ij}(a)\|$ информационную матрицу этой задачи. Тогда

$$I_{ii} = I_{ii}(0) = \int\limits_{-\infty}^{\infty} \frac{|\varphi_i(x)-\varphi_0(x)|^2}{q(x)}\,dx = 2\cdot\frac{M+1}{N},$$

$$I_{ij} = \int\limits_{-\infty}^{\infty} \frac{|\varphi_0(x)|^2}{q(x)}\,dx = \frac{M+1}{N}, \quad i\neq j.$$

Если в качестве функции потерь задачи выбрать функцию $w(x) = \sum\limits_1^N x_i^2 + \Big(\sum\limits_1^N x_i\Big)^2$, то, в силу неравенства Гаека (см. [10, теорема 2.12.1]), для любой оценки $\widehat{a} = (\widehat{a}_1,\dots,\widehat{a}_M)$ параметра a

$$\varliminf_{n\to\infty} \sup_{a\in A} n\mathbf{E}_a W(\widehat{a}-a) \geqslant \mathbf{E}W(\xi), \tag{5.8}$$

где $\xi = (\xi_1,\dots,\xi_M)$ – нормальный случайный вектор со средним нуль и корреляционной матрицей $I_M^{-1} = (\|I_{ij}\|)^{-1}$.

Ортогональным преобразованием матрица I_M приводится к диагональному виду с диагональю $(c(M+1), c, c, \dots, c)$, $c = \frac{M+1}{N}$.

Поэтому, возвращаясь к неравенству (5.7), мы получим (с учетом (5.8)), что для любой оценки $\widehat{\theta}$

$$\begin{aligned}\varliminf_{n} \sup_{\theta\in\Theta} n\,\mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2 &\geqslant \mathbf{E}w(\xi)\\ &= \frac{\sqrt{\det I_M}}{(2\pi)^{M/2}} \int\limits_{R^M} w(x)\exp\Big\{-\frac{1}{2}(I_M x, x)\Big\}\,dx\\ &= \frac{\sqrt{\det I_M}}{(2\pi)^{M/2}} \int\limits_{R^M} \Big(\sum_2^M x_i^2 + (M+1)x_1^2\Big)\\ &\times \exp\Big\{-\frac{1}{2}\Big(c\sum_2^M x_i^2 + c(M+1)x_1^2\Big)\Big\}\,dx = \frac{M}{M+1}N.\end{aligned}$$

Устремляя в этом неравенстве M к бесконечности, мы придем к неравенству теоремы. □

6. Оценка плотности интенсивности пуассоновского случайного множества. В этом разделе мы рассмотрим задачу III в предположении, что неизвестная плотность интенсивности наблюдаемого пуассоновского множества $\theta \in \Theta \subseteq H_K$, где $H_K \subseteq L_2(\mathfrak{X}, \mathfrak{A}, \mu)$ – подпространство с воспроизводящим ядром K. Чтобы избежать ненужных затруднений, мы всегда предполагаем, что множество

$$\mathbf{D} = \{(x, y) : x = y\}$$

измеримо в $(\mathfrak{X}\times\mathfrak{X}, \mathfrak{A}\times\mathfrak{A})$ (см. [11, глава 2]).

Теорема 6.1. *Пусть наблюдаемое пуассоновское случайное множество $\Pi \subset \mathfrak{X}$ имеет плотностью интенсивности (относительно μ) функцию $\theta \in \Theta \subseteq H_K \subseteq L_2$. Оценим θ статистикой*

$$\widehat{\theta}(y) = \sum_{x\in\Pi} K(x, y).$$

Тогда

$$\mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2 = \int\limits_{\mathfrak{X}} K(x, x)\theta(x)\,d\mu. \tag{6.1}$$

В частности,

$$\inf_{\widehat{\theta}} \sup_{\theta} \mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 \leqslant \sup_{\theta} \int\limits_{\mathfrak{X}} K(x,x)\theta(x)\, d\mu. \tag{6.2}$$

Доказательство. В силу теоремы Кэмпбелла (см. [11, глава 3])

$$\mathbf{E}_\theta \widehat{\theta}(y) = \int\limits_{\mathfrak{X}} K(x,y)\theta(x)\, d\mu = (K(\cdot,y),\theta) = \theta(y).$$

Снова по теореме Кэмпбелла

$$\mathbf{D}\widehat{\theta}(y) = \mathbf{E}|\theta(y) - \widehat{\theta}(y)|^2 = \int\limits_{\mathfrak{X}} K^2(x,y)\theta(x)\, d\mu.$$

Интегрируя последнее равенство по мере μ, найдем, что

$$\mathbf{E}_\theta \|\widehat{\theta} - \theta\|^2 = \int\limits_{\mathfrak{X}} \theta(x)\, d\mu(x) \int\limits_{\mathfrak{X}} K^2(x,y)\, d\mu(y) = \int\limits_{\mathfrak{X}} K(x,x)\theta(x)\, d\mu.$$

Теорема доказана. □

Как и в предыдущем случае, полное обращение неравенства (6.2), вообще говоря, невозможно, так как $\theta(x) \geqslant 0$. Мы докажем аналоги теорем 5.2 и 5.3. Чтобы сформулировать эти теоремы, изменим несколько постановку задачи и предположим, что наблюдаемое пуассоновское случайное множество Π_ε имеет плотностью интенсивности функцию $\frac{\theta(x)}{\varepsilon}$, где ε – известный малый параметр, а $\theta \in \Theta$ – функция, подлежащая оценке. Примером подобной постановки служит, например, следующая задача. Наблюдаются n независимых одинаково распределенных пуассоновских множеств $\Pi_1, \Pi_2, \ldots, \Pi_n$ с плотностью интенсивности θ. По этим наблюдениям надлежит оценить θ. Объединение множеств Π_j, $\Pi_\varepsilon = \cup \Pi_j$, есть достаточная статистика задачи, Π_ε – случайное пуассоновское множество интенсивности $n \cdot \theta(x) = \frac{\theta(x)}{\varepsilon}$, $\varepsilon = \frac{1}{n}$. Выбирая в качестве оценки для θ статистику

$$\widehat{\theta}_\varepsilon(y) = \varepsilon \sum_{x \in \Pi_\varepsilon} K(x,y),$$

найдем, повторяя ход доказательства предыдущей теоремы, что

$$\mathbf{E}_\theta \|\widehat{\theta}_\varepsilon - \theta\|^2 = \varepsilon \int\limits_{\mathfrak{X}} K(x,x)\theta(x)\, dx. \tag{6.3}$$

Теорема 6.2. *Пусть наблюдаемый процесс Пуассона* Π_ε *имеет своей плотностью интенсивности функцию* $\varepsilon^{-1}\theta(x)$, *где* $\varepsilon > 0$ – *известный (малый) параметр. Допустим, что* $\theta \in \Theta \subseteq H_K$, *где* H_K *есть пространство с воспроизводящим ядром примера* 2.2, *и что* Θ *содержит все неотрицательные функции* $\theta \in H_K$, *такие что* $\int\limits_{R^d} \theta(x)\,dx \leqslant M$. *Тогда*

$$\varliminf_{\varepsilon\to 0} \inf_{\widehat\theta} \sup_{\theta\in\Theta} \varepsilon^{-1}\mathbf{E}_\theta\|\widehat\theta-\theta\|^2 = \frac{\operatorname{mes}(\mathrm{G})}{(2\pi)^d}\sup_\theta \int\limits_{R^d}\theta(x)\,dx = \frac{\operatorname{mes}(\mathrm{G})}{(2\pi)^d}\cdot M. \quad (6.4)$$

Доказательство. В силу (6.3) для оценки

$$\widehat\theta_\varepsilon(y) = \varepsilon\sum_{x\in\Pi_\varepsilon} K(x,y)$$

выполняется неравенство

$$\mathbf{E}_\theta\|\widehat\theta_\varepsilon-\theta\|^2 \leqslant \varepsilon\int\limits_{R^d} K(x,x)\theta(x)\,dx = \frac{\operatorname{mes}(\mathrm{G})}{(2\pi)^d}\int\limits_{R^d}\theta(x)\,dx,$$

так что заведомо

$$\lim_{\varepsilon\to 0}\inf_{\widehat\theta}\sup_\theta \varepsilon^{-1}\mathbf{E}_\theta\|\widehat\theta-\theta\|^2 \leqslant \frac{\operatorname{mes}(\mathrm{G})}{(2\pi)^d}M.$$

Перейдем к доказательству обратного неравенства. При этом без потери общности можно считать $\int\limits_{R^d}\theta(x)\,dx \leqslant 1$, что и предполагается ниже. Доказательство сходно с доказательством теоремы 5.2, т.е. теоремы 1 из [6], и мы позволим себе опускать некоторые детали доказательства.

Положим

$$\Delta_2^2 = \inf_{\widehat\theta}\sup_{\theta\in\Theta}\mathbf{E}_\theta\|\widehat\theta-\theta\|^2$$

и оценим снизу величину $\varepsilon^{-1}\Delta_2^2$ (следует иметь в виду, что $\mathbf{E}_\theta(\cdot)$ означает математическое ожидание по отношению к пуассоновскому полю Π_ε с плотностью интенсивности $\theta\varepsilon^{-1}$ (а не θ)).

Зафиксируем малое положительное число $\delta > 0$ и рассмотрим множество $G_\delta = (1-\delta)G = \{y : y = (1-\delta)x, x \in G\}$. Обозначим Γ_γ куб $\{x : |x_i| \leqslant \gamma\}$ и выберем $\gamma = \gamma(\delta)$ так, чтобы $G_\delta + \Gamma_\gamma = \{x : x = x^1 + x^2, x^1 \in K_\delta, x^2 \in \Gamma_\gamma\} \subseteq G$. Условимся обозначать $\mathcal{E}(D)$ множество функций $f \in L_2(R^d)$, преобразование Фурье которых обращается в нуль вне множества D.

Положим

$$q_\gamma(x) = \prod_1^d \frac{\sin^2(x_i\gamma/2)}{\pi\gamma x_i^2}.$$

Очевидно, что $q_\gamma \in \mathcal{E}(\Gamma_\gamma)$ и если $\psi \in \mathcal{E}(G_\delta)$, то $\psi q_\gamma \in \mathcal{E}(G)$. Ниже для краткости вместо q_γ будем писать просто q. Пусть функции $\{\varphi_j\}$ образуют полную ортонормальную систему в $\mathcal{E}(G_\delta)$. Рассмотрим семейство функций

$$\theta(x;a) = q(x)\biggl(1 + \sum_1^N a_j\varphi_j(x)\biggr),$$

зависящее от N-мерного параметра $a = (a_1, \dots, a_N)$. Все эти функции суть элементы H_K. Обозначая $\widetilde{\varphi}_j$ преобразование Фурье функции φ_j, найдем

$$\|\varphi_j\| \leqslant (2\pi)^{-d} \int\limits_{G_\delta} |\widetilde{\varphi}_j(t)|\,dt < (\operatorname{mes} \mathrm{G})^{1/2}.$$

Поэтому если точка $a = (a_1, \dots, a_N)$ принадлежит кубу

$$A = \biggl\{a : |a_j| \leqslant \frac{\sqrt{\operatorname{mes} \mathrm{G}}}{N}\biggr\},$$

то все функции $\theta(x;a)$ суть плотности интенсивности из Θ. Условимся обозначать через A также и множество функций $\theta(x;a)$, $a \in A$.

Таким образом, $A \subseteq \Theta$ и для любой оценки $\widehat{\theta}_\varepsilon$ плотности θ

$$\sup_{\theta\in\Theta} \mathbf{E}_\theta\|\widehat{\theta}_\varepsilon - \theta\|^2 \geqslant \sup_{\theta\in A} \mathbf{E}_\theta\|\widehat{\theta}_\varepsilon - \theta\|^2. \tag{6.5}$$

Оценивая функцию $\theta \in A$, мы заведомо можем ограничиться оценками $\widehat{\theta}_\varepsilon$ вида

$$\widehat{\theta}_\varepsilon(x;\Pi_\varepsilon) = q(x)\biggl(\widehat{a}_0 + \sum_1^N \widehat{a}_j\varphi_j(x)\biggr), \quad \widehat{a}_j = \widehat{a}_j(\Pi_\varepsilon).$$

Поэтому

$$\sup_{\theta\in A} \mathbf{E}_\theta\|\widehat{\theta} - \theta\|^2 = \sup_{\theta\in A} \frac{1}{\varepsilon}\mathbf{E}_\theta\biggl\|q(\widehat{a}_0 - 1) + \sum_1^N (\widehat{a}_j - a_j)\varphi_j\biggr\|^2. \tag{6.6}$$

От компоненты $\widehat{a}_0$ можно избавиться следующим образом.

Рассмотрим матрицу

$$B_N = \left\| \int\limits_{R^d} \varphi_j(x)\varphi_i(x)q^2(x)\,dx \right\|$$

и запишем неотрицательную квадратичную форму переменных t, $z = (z_1, \dots, z_N)$

$$Q = \left\| q\left(t + \sum_1^N z_j\varphi_j\right)\right\|^2$$

в виде

$$Q = \Big(B_N\big(z + B_N^{-1}u\big), \big(z + B_N^{-1}u\big)\Big) + \Big[t\|q\|^2 - \big(B_N^{-1}u, u\big)\Big], \tag{6.7}$$

где $u = (u_1, \dots, u_N)$ – вектор с компонентами $t \int\limits_{R^d} q^2(x)\varphi_j(x)\,dx$. Форма Q – неотрицательная, и, полагая в (6.7) $z = -B^{-1}u$, мы найдем, что $t\|q\|^2 - (B_N^{-1}u, u) \geqslant 0$ и что, следовательно,

$$Q \geqslant \Big(B_N\big(z + B_N^{-1}u\big), \big(z + B_N^{-1}u\big)\Big). \tag{6.8}$$

Заменим теперь в (6.6) оценку $(\widehat{a}_0, \widehat{a}_1, \dots, \widehat{a}_N)$ оценкой $(\widetilde{a}_1, \dots, \widetilde{a}_N)$, где $\widetilde{a}_j = \widehat{a}_j - B_N^{-1}u$. Мы получим тогда, что

$$\inf_{\{a_j\}} \sup_{\theta\in A} \mathbf{E}_\theta \left\| q(\widehat{a}_0 - 1) + \sum_1^N (\widehat{a}_j - a_j)\varphi_j \right\|^2 \geqslant \inf_{\{a_j\}} \sup_{\theta\in A} \mathbf{E}_\theta \left\| q\sum_1^N (\widehat{a}_j - a_j)\varphi_j \right\|^2$$

и что, следовательно,

$$\inf_{\widehat{\theta}_\varepsilon} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}_\varepsilon - \theta\|^2 \geqslant \inf_{\{a_j\}} \sup_{\theta\in A} \mathbf{E}_\theta \left\| q\sum_1^N (\widehat{a}_j - a_j)\varphi_j \right\|^2. \tag{6.9}$$

Чтобы оценить правую часть последнего неравенства, рассмотрим следующую вспомогательную задачу параметрического оценивания: по наблюдению пуассоновского случайного множества Π_ε с плотностью интенсивности $\varepsilon^{-1}\theta(x; a)$ оценить параметр $a = (a_1, \dots, a_N) \in A$ в предположении, что функция потерь $w(z) = \left\| q\sum_1^N z_j\varphi_j \right\|^2$.

Если $\mathbf{P}_\lambda$ есть распределение пуассоновского множества Π на $(\mathfrak{X}, \mathfrak{A}, \mu)$ с плотностью интенсивности λ, то

$$\frac{d\mathbf{P}_{\lambda_2}}{d\mathbf{P}_{\lambda_1}}(\Pi) = \exp\left\{ \sum_{x\in\Pi} \ln\frac{\lambda_2(x)}{\lambda_1(x)} - \int\limits_{\mathfrak{X}} [\lambda_2(x) - \lambda_1(x)]\,d\mu \right\}$$

(см., например, [12, с. 124], или [13]).

Отсюда легко следует, что количество информации Фишера параметрического статистического эксперимента порожденного Π_ε определено и информационная матрица Фишера в точке $a = 0$

$$I_N = \left\| \int\limits_{R^d} \varphi_i(x)\varphi_j(x)q(x)\,dx \right\|$$

и что соответствующий статистический эксперимент удовлетворяет условию локальной асимптотической нормальности (ЛАН) с нормирующим множителем $\varphi(\varepsilon) = \sqrt{\varepsilon}\, I_N^{-1}$, $\varepsilon \to 0$ (см. [10, глава II]). Об условиях ЛАН для пуассоновских наблюдений см. также [14].

В силу теоремы Гаека (см. [10, §2.12]) для любой оценки $\widehat{a}$ параметра a

$$\varliminf_{\varepsilon\to 0} \sup_{a\in A} \mathbf{E}_a^{(\varepsilon)} W\left(\varepsilon^{-\frac{1}{2}}(\widehat{a} - a)\right) \geqslant \mathbf{E}W(\xi),$$

где ξ – нормальный случайный вектор с нулевым средним и корреляционной матрицей I_N^{-1}. Последнее неравенство можно переписать в виде

$$\begin{aligned}
&\varliminf_{\varepsilon\to 0} \sup_{a\in A} \frac{1}{\varepsilon}\mathbf{E}_a \left\| q\sum_1^N (\widehat{a}_j - a_j)\varphi_j \right\|^2 \\
&\qquad \geqslant \frac{\sqrt{\det I_N}}{(2\pi)^{N/2}} \int\limits_{R^N} dv \left\{ \int\limits_{R^d} q^2(y) \left| \sum_1^N v_j\varphi_j(y) \right|^2 dy \right\} \\
&\qquad\qquad \times \exp\left\{ -\frac{1}{2}\int\limits_{R^d} q(x) \left\{ \int\limits_{R^d} q^2(y) \left| \sum_1^N v_j\varphi_j(y) \right|^2 dx \right\} \right\}
\end{aligned}$$

Выражение справа можно записать в виде $\mathbf{E}(B_N\xi, \xi)$, где матрица

$$B_N = \left\| \int\limits_{R^d} q^2(x)\varphi_i(x)\varphi_j(x)\,dx \right\|,$$

а $\xi = (\xi_1 \ldots, \xi_N)$ – нормальный случайный вектор со средним нуль и корреляционной матрицей I_N^{-1}. Следуя [6] (см. соотношение (17) в [6]), нетрудно показать, что

$$(B_N\xi, \xi) \geqslant (I_N^2\xi, \xi) = |I_N\xi|^2.$$

Далее,

$$\mathbf{E}|I_N\xi|^2 = \operatorname{tr} I_N = \int\limits_{R^k} q(x)\left(\sum_1^N \varphi_j^2(x)\right) dx,$$

так что для любого N

$$\varliminf_{\varepsilon\to 0} \inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \geqslant \int\limits_{R^d} \sum_1^N \varphi_j^2(x) q(x)\, dx.$$

Устремляя здесь $N \to \infty$ и учитывая, что $\sum\limits_1^\infty \varphi_j^2(x) = \dfrac{\operatorname{mes} \mathrm{G}_\delta}{(2\pi)^d}$ (см. пример 2.2), мы найдем, что

$$\varliminf_{\varepsilon\to 0} \inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \geqslant \frac{\operatorname{mes} \mathrm{G}_\delta}{(2\pi)^d} \underset{\delta\to 0}{\to} \frac{\operatorname{mes} \mathrm{G}}{(2\pi)^d}.$$

Теорема доказана. □

Приведем еще аналог теоремы 5.3. Рассмотрим задачу III оценки плотности интенсивности $\theta(x)$ по наблюдению пуассоновского случайного множества Π_ε с интенсивностью $\frac{\theta(x)}{\varepsilon}$ в предположении, что $\theta \in \Theta$, $\Theta \subseteq H_K$ и содержит все плотности интенсивности $\theta(x)$ с $\int\limits_{-\infty}^{\infty} \theta(x)\, dx \leqslant M < \infty$, а H_K есть пространство с воспроизводящим ядром примера 2.3, причем $a_j = \frac{j}{N}$, $j = \cdots -1, 0, 1, \ldots$, а $\varphi_j(x) = \sqrt{N}$ для $x \in [a_j, a_{j+1})$. Тогда

$$\varliminf_{\varepsilon\to 0} \inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \varepsilon^{-1}\, \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 = \sup_{\theta\in\Theta} \int\limits_{-\infty}^{\infty} K(x,x)\theta(x)\, dx = MN.$$

Доказательство. Неравенство

$$\inf_{\widehat{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}-\theta\|^2 \leqslant \varepsilon \sup_{\theta\in\Theta} \int\limits_{-\infty}^{\infty} K(x,x)\theta(x)\, dx = MN\varepsilon$$

следует из (6.3).

Докажем обратное неравенство. Достаточно ограничиться случаем $M = 1$. Пусть $\delta > 0$ – малое положительное число. Рассмотрим

семейство функций

$$\theta(x;a) = \sqrt{N}\ (1-\delta)a\varphi_1(x),$$

зависящее от вещественного параметра a, $a \in A = \{a : |1-a| \leqslant \delta\}$. Если $\delta \leqslant \frac{1}{2}$, то для всех $a \in A$ функции $\theta(x;a) \in \Theta$. Будем обозначать это семейство θ тоже буквой A. Тогда

$$\sup_{\theta\in\Theta}\ \mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2 \geqslant \sup_{\theta\in A}\ \mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2. \tag{6.12}$$

Очевидно, что оценивая $\theta \in \Theta$ мы можем ограничиться оценками вида $\widehat{\theta} = \sqrt{N}\,(1-\delta)\widehat{a}\varphi_1$, так что

$$\inf_{\widehat{\theta}}\ \sup_{\theta\in A}\ \mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2 = N(1-\delta)^2\ \inf_{\widehat{a}}\ \sup_{\theta\in A}\ \mathbf{E}_\theta|\widehat{a}-a|^2. \tag{6.13}$$

Чтобы оценить правую часть последнего неравенства, рассмотрим задачу оценивания вещественного параметра $a \in A$ по наблюдениям пуассоновского случайного множества с интенсивностью $\varepsilon^{-1}\theta(x;a)$. Соответствующие статистические эксперименты удовлетворяют условию ЛАН и имеют информацию Фишера

$$I(a) = \frac{1}{\varepsilon}\ \frac{1}{a(1-\delta)}.$$

В силу неравенства Гаека (см. [10, теорема 2.12.1]

$$\lim_{\varepsilon\to 0}\ \inf_{\widehat{a}}\ \sup_{\theta\in A}\ \varepsilon^{-1}\,\mathbf{E}_\theta|\widehat{a}-a|^2 \geqslant \frac{1}{I(a)} = \frac{1}{1-\delta}.$$

Отсюда и из неравенств (6.12), (6.13) следует, что

$$\lim_{\varepsilon\to 0}\inf_{\widehat{\theta}}\ \sup_{\theta\in\Theta}\ \varepsilon^{-1}\,\mathbf{E}_\theta\|\widehat{\theta}-\theta\|^2 \geqslant (1-\delta)N.$$

Устремляя δ к нулю, придем к утверждению теоремы. □

Замечание. Любопытно, что верхняя грань достигается уже на плотностях из одномерного подпространства H_K, натянутого на φ_1, ср. с доказательством теоремы 5.2, где используются все подпространства, натянутые на $\varphi_1, \ldots, \varphi_M$ для всех M.

7. Оценка производных функции θ. Здесь мы лишь слегка коснемся этой задачи (см., например, [15]), ограничившись для определенности задачей оценки производных плотности распределения.

Итак, допустим, что наблюдается выборка

$$X_1, \dots, X_n,$$

состоящая из n независимых одинаково распределенных случайных величин x_j, принимающих значения в R^d и имеющих там плотность распределения $\theta(x)$. Предполагается, что $\theta \in \Theta$, где Θ – заданное подмножество $L_2(R^d)$. Задача заключается в оценке производных $\theta^{(\overline{\nu})}(x) = \frac{\partial^{\nu_1}}{\partial x_1^{\nu_1}} \cdots \frac{\partial^{\nu_d}}{\partial x_d^{\nu_d}} \theta(x)$ в предположении, что $\Theta \subseteq H_K$, $H_K \subseteq L_2$ – подпространство с воспроизводящим ядром K. Допустим, что если $f \in H_K$, то и все произведные f определенного порядка принадлежат H_K. Для простоты мы даже предположим, что f из H_K бесконечно дифференцируемы и все производные лежат в H_K.

Рассмотрим в качестве оценки для $\theta^{(\overline{\nu})}$ статистику

$$\widehat{\theta}_n(y) = \frac{1}{n} \sum K_y^{(\overline{\nu})}(X_j, y).$$

Не вдаваясь в детали, когда приводимые ниже выкладки обоснованы, найдем

$$\mathbf{E}_\theta \, \widehat{\theta}_n(y) = \int\limits_{R^d} K_y^{(\overline{\nu})}(x, y)\theta(x)\, dx = \left(\int\limits_{R^d} K(x, y)\theta(x)\, dx \right)_y^{(\overline{\nu})} = \theta^{(\overline{\nu})}(y),$$

так что $\widehat{\theta}_n$ – несмещенная оценка. Далее (ср. с п. 4)

$$\begin{aligned} \mathbf{E}_\theta \|\widehat{\theta}_n - \theta^{(\overline{\nu})}\|^2 &\leqslant \frac{1}{n} \int\limits_{R^d} dy \int \left| K_y^{(\overline{\nu})}(x, y) \right|^2 \theta(x)\, dx \\ &= \frac{1}{n} \int\limits_{R^d} \theta(x)\, dx \int\limits_{R^d} K_y^{(\overline{\nu})}(x, y) K_y^{(\overline{\nu})}(x, y)\, dy. \end{aligned}$$

Интегрирование по частям превращает внутренний интеграл в $\left[K_y^{(2\overline{\nu})}(x, y)\right]_{y=x}$, и мы приходим к следующему результату:

для оценки $\widehat{\theta}_n$ справедливо неравенство

$$\mathbf{E}_\theta \|\widehat{\theta}_n - \theta^{(\overline{\nu})}\|^2 \leqslant \frac{1}{n} \int\limits_{R^d} \left[K_y^{(2\overline{\nu})}(x, y) \right]_{y=x} \theta(x)\, dx.$$

Пример 7.1. Пусть H_K – пространство функций, преобразование Фурье которых обращается в нуль вне ограниченного множества G положительной меры. H_K есть пространство целых функций f экспоненциального типа и вместе с f содержит все производные $f^{(\overline{\nu})}$. В данном случае все приведенные выше выкладки, очевидно, законны и для $\Theta \subseteq H_K$

$$\varliminf_{n\to\infty} \inf_{\widehat{\theta}_n} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}_n - \theta^{(\nu)}\|^2 \leqslant (2\pi)^{-d} \int\limits_G x_1^{2\nu_1}, \ldots, x_d^{2\nu_d}\, dx.$$

Можно показать, что если Θ содержит все плотности из H_K, то это неравенство правращается в равенство (ср. с теоремой 4.1).

8. В этом пункте мы покажем, как осуществляется вторая часть метода Н. Н. Ченцова, связанная с аппроксимацией множеств $\Theta \subseteq L_2(\mathfrak{X}, \mathfrak{A}, \mu)$ пространствами H_K с воспроизводящим ядром K.

Следуя [7], обозначим $\mathcal{K}(T,p)$, $1 \leqslant p \leqslant \infty$, совокупность всех воспроизводящих ядер в L_2, таких, что $K(x,x) \in L_p(\mathfrak{X}, \mathfrak{A}, \mu)$ и

$$\|K(x,x)\|_p \leqslant T. \tag{8.1}$$

Для множества $\Theta \subseteq L_2$ определим поперечники $\delta_T(\Theta; p)$ равенством

$$\delta_T(\Theta; p) = \inf_{a,H} \sup_{\theta\in\Theta} \inf_{h\in H+a} \|\theta - h\|, \tag{8.2}$$

где внешний inf берется по всем $H = H_K$ с $K \in \mathcal{K}(T,p)$ и всем векторам $a \in L_2$.

Пример 8.1. Всякое N-мерное подпространство H с ортонормальным базисом $\{\varphi_1, \ldots, \varphi_N\}$ имеет воспроизводящее ядро $K \in \mathcal{K}(N,1)$. Поэтому $\delta_N(\Theta, 1) \leqslant d_N(\Theta)$, где d_N – N-й поперечник Колмогорова множества Θ.

Схема построения оценок функционального параметра $\theta \in \Theta$ выглядит следующим образом.

Рассмотрим для определенности задачу оценивания функции $\theta \in \Theta$ по наблюдению пуассоновского случайного множества Π_ε в $\mathfrak{X}$ с плотностью интенсивности $\varepsilon^{-1}\theta(x)$. Относительно Θ будем дополнительно предполагать, что $\sup\limits_{\theta\in\Theta} \|\theta\|_1 \leqslant A$. Положим

$$\delta_T = \delta_T(\Theta) = \delta_T(\Theta, \infty).$$

Теорема 8.1 (ср. [7, теорема 1]). *Имеет место неравенство*

$$\inf_{\widehat{\theta}_\varepsilon} \sup_{\theta\in\Theta} \mathbf{E}_\theta \|\widehat{\theta}_\varepsilon - \theta\|^2 \leqslant \inf_T (4\delta_T^2 + \varepsilon T A). \tag{8.3}$$

Доказательство. Будем для простоты считать, что при определении $\delta_T(\theta)$ сдвиг $a = 0$. Подберем такое H_K, что для него расстояние между H_K и Θ не превосходит $2\delta_T$. Выберем в качестве оценки θ статистику

$$\widehat{\theta}_K(y) = \varepsilon \sum_{x\in\Pi_\varepsilon} K(x,y).$$

Пусть $\theta_K(y) = \int\limits_{\mathfrak{X}} K(x,y)\theta(x)\,dx$ – проекция θ на H_K. В силу результатов раздела 6, теорема 6.1, $\widehat{\theta}_K$ есть несмещенная оценка θ_K, $\mathbf{E}_\theta \widehat{\theta}_K(y) = \theta_K(y)$. Как мы видели выше, теорема 6.1,

$$\mathbf{E}_\theta \|\widehat{\theta}_K - \theta_K\|^2 = \varepsilon \int\limits_{\mathfrak{X}} K(x,x)\theta(x)\,dx.$$

Поэтому

$$\begin{aligned}\mathbf{E}_\theta \|\theta - \widehat{\theta}_K\|^2 &= \mathbf{E}_\theta \Big\{ \|\widehat{\theta}_K - \theta_K\|^2 \Big\} + \|\theta - \theta_K\|^2 \\ &\leqslant \varepsilon \int\limits_{\mathfrak{X}} K(x,x)\theta(x)\,dx + 4\delta_T^2 \leqslant \varepsilon T \sup_{\theta\in\Theta} \int\limits_{\mathfrak{X}} \theta(x)\,dx + 4\delta_T^2.\end{aligned}$$

Теорема доказана. □

Рассмотрим два примера применения этой теоремы.

Пример 8.2. Рассмотрим задачу оценивания функции θ по наблюдению пуассоновского множества $\Pi_\varepsilon \subseteq R^d$ с плотностью интенсивности $\varepsilon^{-1}\theta(x)$. Неизвестная функция $\theta \in \Theta \subseteq L_2(R^d)$, а Θ – известный класс функций.

Пусть $\varphi(x)$, $x \in R^d$ – непрерывная положительная функция, такая что $\varphi(x) \to \infty$, когда $|x| \to \infty$. Обозначим $F(\varphi; L)$ класс функций $f \in L_2(R^d)$, таких что

$$\int\limits_{R^d} \varphi^2(t)|g(t)|^2 \leqslant L^2, \tag{8.4}$$

где $g(t) = \int\limits_{R^d} e^{i(t,x)} f(x)\,dx$ – преобразование Фурье функции $f(x)$. Условие (8.4) можно рассматривать как некоторое условие гладкости функций из $F(\varphi; L)$. Классы $F(\varphi; L)$ можно также трактовать как континуальные аналоги эллипсоидов.

Теорема 8.1. *Пусть* $\Theta \subseteq F(\varphi; L)$ *и пусть* $\int\limits_{R^d} \theta(x)\,dx \leqslant 1$ *для* $\theta \in \Theta$. *Пусть функция*

$$a(T) = \int\limits_{\{t:|\varphi(t)|\leqslant T\}} dt.$$

Существует оценка $\widehat{\theta}_\varepsilon$ *функции* θ, *такая что*

$$\mathbf{E}_\theta \|\widehat{\theta}_\varepsilon - \theta\| \leqslant \inf_T \left(\frac{L^2}{T^2} + \varepsilon \frac{a(T)}{(2\pi)^d} \right). \tag{8.5}$$

Доказательство. Обозначим L_T подпространство $L_2(R^d)$, состоящее из всех функций $f \in L_2(R^d)$, преобразование Фурье g которых обращается в нуль вне множества $G = G_T = \{t : \varphi(t) \leqslant T\}$. Мы видели (пример 2.1), что L_T есть пространство H_K с воспроизводящим ядром

$$K(x, y) = \frac{1}{(2\pi)^d} \int\limits_G e^{-i(t,x-y)}\,dt.$$

Обозначим θ_T проекцию θ на $L_T = H_K$ и рассмотрим оценку

$$\widehat{\theta}(y) = \varepsilon \sum_{x \in \Pi_\varepsilon} K(x, y).$$

Для этой оценки

$$\mathbf{E}_\theta \|\widehat{\theta}_T - \theta\|^2 = \|\theta - \theta_T\|^2 + \mathbf{E}_\theta \|\widehat{\theta}_T - \theta_T\|^2. \tag{8.6}$$

Обозначая g преобразование Фурье функции θ, найдем

$$\|\theta - \theta_T\|^2 = \int\limits_{\{t:\varphi(t)>T\}} |g(t)|^2\,dt \leqslant \frac{1}{T^2} \int\limits_{R^d} \varphi^2(t)|g(t)|^2\,dt \leqslant \frac{L^2}{T^2}.$$

Второе слагаемое в (8.6)

$$\mathbf{E}_\theta\|\widehat{\theta}_T-\theta_T\|^2\leqslant\varepsilon\int\limits_{R^d}K(x,x)\theta(x)\,dx=\varepsilon\frac{a(T)}{(2\pi)^d}.$$

Таким образом,

$$\mathbf{E}_\theta\|\widehat{\theta}_T-\theta_T\|^2\leqslant\frac{L^2}{T^2}+\varepsilon\frac{a(T)}{(2\pi)^d}. \tag{8.7}$$

Обозначая $\widehat{\theta}_\varepsilon$ оценку, соответствующую $\widehat{\theta}_T$ с T, минимизирующим правую часть (8.7), придем к неравенству теоремы. □

Пусть вектор $\overline{\beta}=(\beta_1,\dots,\beta_d)$, $\beta_j>0$. Положим

$$E^{(1)}_{\overline{\beta}}(L)=F(\varphi_1;L),\quad \varphi_1(x)=\Big(\big(1+|x_1|\big)^{2\beta_1}+\dots+\big(1+|x_d|\big)^{2\beta_d}\Big)^{1/2},$$
$$E^{(2)}_{\overline{\beta}}(L)=F(\varphi_2;L),\quad \varphi_2(x)=\Big(\big|x_1\big|^{2\beta_1}+\dots+\big|x_d\big|^{2\beta_d}\Big)^{1/2}.$$

Теорема 8.3. *Пусть* $\Theta\subseteq E^{(1)}_{\overline{\beta}}(L)$. *Допустим, что* $\min\beta_j=\beta=\beta_1=\dots=\beta_l<\beta_{l+1}\leqslant\dots\leqslant\beta_d$. *Тогда существует такая оценка* $\widehat{\theta}_\varepsilon$ *функции* θ, *что*

$$\mathbf{E}^{(\varepsilon)}_\theta\|\widehat{\theta}_\varepsilon-\theta\|^2\leqslant c\Big(\varepsilon\Big(\ln\frac{1}{\varepsilon}\Big)^{l-1}\Big)^{\frac{2\beta}{2\beta+1}}. \tag{8.8}$$

Константа c *зависит от* L *и* $\overline{\beta}$.

Доказательство. Когда $T\to\infty$,

$$a(T)=\int\limits_{\{x:\varphi_1(x)\leqslant T\}}dx\sim c_0T^{1/\beta}(\ln T)^{l-1}$$

и (8.8) следует из (8.5). □

Теорема 8.4. *Пусть* $\Theta\subseteq E^{(2)}_{\overline{\beta}}(L)$. *Определим* β *соотношением* $\frac{1}{\beta}=\frac{1}{\beta_1}+\cdot+\frac{1}{\beta_d}$. *Существует оценка* $\widehat{\theta}_\varepsilon$ *функции* θ, *такая что*

$$\mathbf{E}^{(\varepsilon)}_\theta\|\widehat{\theta}_\varepsilon-\theta\|^2\leqslant c\varepsilon^{\frac{2\beta}{2\beta+1}}. \tag{8.9}$$

Постоянная c *зависит от* L *и* $\overline{\beta}$.

Доказательство. На этот раз

$$a(T) = \int\limits_{\{x:\varphi(x)\leqslant T\}} dx = T^{1/\beta} \int\limits_{\{x:\varphi(x)\leqslant 1\}} dx,$$

и (8.9) следует из (8.5). □

Пример 8.3. В следующем примере мы снова рассмотрим задачу оценивания функции $\theta(x)$ по наблюдениям случайного пуассоновского множества $\Pi_\varepsilon \subseteq R^1$ с плотностью интенсивности $\varepsilon^{-1}\theta(x)$. Неизвестная функция $\theta \in \Theta \subseteq L_2(R^1)$. Задавшись числом $T > 0$, введем пространство с вопроизводящим ядром H_K примера 2.3, полагая $a_j = \frac{j}{T}$, а $\varphi_j(x) = \sqrt{T}$ для $a_j \leqslant x < a_{j+1}$. Пусть $K(x,y) = \sum \varphi_j(x)\varphi_j(y)$ – воспроизводящее ядро пространства H_K. Обозначим θ_T проекцию θ на H_K. В качестве оценки θ рассмотрим статистику

$$\widehat{\theta}_T(y) = \varepsilon \sum_{x\in\Pi_\varepsilon} K(x,y).$$

Эта статистика есть классическая гистограмма плотности интенсивности с ячейками ширины T^{-1}. Как мы видели, $\widehat{\theta}_T$ – несмещенная оценка проекции θ_T, $\mathbf{E}\widehat{\theta}_T = \theta_T$, и

$$\mathbf{E}_\theta \|\theta - \widehat{\theta}_T\|^2 \leqslant \varepsilon \int\limits_{-\infty}^{\infty} K(x,x)\theta(x)\,dx + \|\theta - \theta_T\|^2 = T\varepsilon \cdot \int\limits_{-\infty}^{\infty} \theta(x)\,dx + \|\theta - \theta_T\|^2.$$

Обозначим $\omega(h;\theta)$ модуль непрерывности θ в $L_2(R)$. Имеем

$$\|\theta - \theta_T\|^2 = T \sum_\nu \int\limits_{\nu/T}^{(\nu+1)/T} \int\limits_{\nu/T}^{(\nu+1)/T} |\theta(x) - \theta(y)|^2\,dx\,dy \leqslant \omega\Big(\frac{1}{T};\theta\Big).$$

Из двух последних неравенств следует, что

$$\mathbf{E}_\theta \|\theta - \widehat{\theta}_T\|^2 \leqslant T\varepsilon \sup_\theta \int\limits_{-\infty}^{\infty} \theta(x)\,dx + \sup_\theta \omega^2\Big(\frac{1}{T};\theta\Big).$$

Таким образом, существует оценка θ_ε функции θ, такая что

$$\mathbf{E}_\theta \|\theta - \widehat{\theta}_T\|^2 \leqslant \inf_T \left(\varepsilon T \sup_\theta \int\limits_{-\infty}^{\infty} \theta(x)\, dx + \sup_\theta \omega^2\Big(\frac{1}{T}; \theta\Big) \right).$$

Если, например, Θ состоит из функций θ, для которых $\int\limits_{-\infty}^{\infty} \theta(x) dx \leqslant 1$, а $\omega(h;0) \leqslant Lh^\beta$, то найдется оценка $\widehat{\theta}_\varepsilon$, такая что

$$\mathbf{E}_\theta \|\widehat{\theta}_\varepsilon - \theta\|^2 \leqslant c\varepsilon^{\frac{2\beta}{2\beta+1}},$$

где постоянная c зависит от β, L.

Последнее неравенство, как и неравенства предыдущего примера, относящиеся к классам $E^{(i)}$, с точностью до константы неулучшаемы.

Литература

1. Н. Н. Ченцов, *Оценка неизвестной плотности распределения по наблюдениям.* — ДАН СССР **147**, No. 1 (1962), 45–48.
2. А. Н. Колмогоров, *Über die beste Annäherung von Funktionen einer gegebenen Funktionalklassen.* — Ann. Math., **37**, No. 1 (1936), 107–111; (см. также А. Н. Колмогоров, *Избранные труды* — Математика и механика, Наука, М., 1985).
3. В. М. Тихомиров, *Некоторые вопросы теории приближений.* Изд-во МГУ, М., 1976.
4. В. М. Тихомиров, *Теория приближений.* — Итоги науки и техники. Современные проблемы математики **14** (1987), 103–270.
5. К. И. Бабенко (ред.), *Теоретические основы и конструирование численных алгоритмов задач математической физики.* Наука, М., 1979.
6. И. А. Ибрагимов, Р. З. Хасьминский, *Об оценке плотности распределения, принадлежащей одному классу целых функций.* — Теория вероятн. и ее примен. **27**, вып. 3 (1982), 514–524.
7. И. А. Ибрагимов, *Об одной характеристике точности оценивания плотности распределения.* — Теория вероятн. и ее примен. **38**, вып. 2 (1993), 425–431.
8. К. Иосида, *Функциональный анализ.* Мир, М., 1967.
9. N. Aronzain, *Theory of reproducing kernels.* — Trans. Amer. Math. Soc. **68** (1950), 337–404.
10. И. А. Ибрагимов, Р. З. Хасьминский, *Асимптотическая теория оценивая.* Наука, М., 1979.
11. Дж. Кингман, *Пуассоновские процессы.* МЦНМО, М., 2007.
12. U. Grenander, *Abstract Inference.* J. Wiley, New York, 1981.
13. M. Brown, *Discrimination of Poisson processes.* — Ann. Math. Statist. **42** (1971), 773–776.
14. Yu. A. Kutoyants, *Introduction to Statistics of Poisson Processes.* Springer, 2012.

15. И. А. Ибрагимов, Р. З. Хасьминский, *Об оценках сигнала, его производных и точки максимума для гауссовских наблюдений.* — Теория вероятн. и ее примен. **25**, вып. 4 (1980), 718–733.

Ibragimov I. A. A generalization of Chentsov's projection estimates.

In 1962, N. N. Chentsov suggested the following method of estimation a functional parameter θ belonging to a Hilbert space H. He suggested to project θ on finite-dimensional subspaces of H and consider as estimates of θ estimates of these projections. In this paper, we suggest to consider the projections on all reproducing kernel subspaces of H.

Санкт-Петербургское отделение
Математического института
им. В. А. Стеклова РАН,
Фонтанка 27, 191023
и Санкт-Петербургский
государственный университет,
Математико–механический факультет
Университетский пр., 28, Петродворец,
198504 Санкт-Петербург, Россия
E-mail: ibr32@pdmi.ru

Поступило 23 апреля 2013 г.