Записки научных
семинаров ПОМИ
Том 368, 2009 г.

Ю. И. Ингстер, И. А. Суслина

АДАПТИВНОЕ ОБНАРУЖЕНИЕ ФУНКЦИЙ БОЛЬШОГО ЧИСЛА ПЕРЕМЕННЫХ

1. Введение

В этой работе мы будем рассматривать наблюдения вида

$$X = f + \varepsilon N, \tag{1.1}$$

где $f \in L_2([0,1]^d) \stackrel{\Delta}{=} L_2^d$, а N – гауссовский белый шум на единичном кубе размерности d, $1 \le d \le \infty$. Напомним, что в этой модели наблюдение есть функция $X : L_2^d \to G$, принимающая значения в множестве G гауссовских случайных величин и, если $\xi = X(\phi)$, $\eta = X(\psi)$, $\phi, \psi \in L_2^d$, то $E(\xi) = (f, \phi)$, $E(\eta) = (f, \psi)$, $\mathrm{Cov}\,(\xi, \eta) = \varepsilon^2(\phi, \psi)$. Для любой функции $f \in L_2^d$ наблюдение (1.1) определяет гауссовскую меру $P_{\varepsilon,f}$. См., например, [13, 2].

Фиксируем ортонормальный базис $\{\phi_l,\, l \in \mathcal{L}\}$ в L_2^d, где $\mathcal{L}$ – счетное множество. Можно рассмотреть эквивалентную гауссовскую модель в пространстве последовательностей

$$X_l = \theta_l + \varepsilon\xi_l, \quad \xi_l \sim \mathcal{N}(0,1) \quad \text{i.i.d.}, \quad l \in \mathcal{L},$$

где $\theta_l = (f, \phi_l)$ – коэффициенты ряда Фурье неизвестной функции f, а $X_l = X(\phi_l)$ – эмпирические коэффициенты Фурье.

Нас интересует задача обнаружения f, то есть мы хотим проверить гипотезу $H_0 : f = 0$. Хорошо известно, что при минимаксном подходе задача не имеет удовлетворительного решения без некоторых условий регулярности, наложенных на класс функций f. Обычно ограничения имеют вид $f \in \mathcal{F}$, где множество $\mathcal{F} \subset L_2([0,1]^d)$ состоит из “достаточно регулярных” функций. Часто ограничения регулярности описываются в терминах некоторой нормы, характеризующей свойства гладкости рассматриваемых функций (нормы Соболева, Гельдера, Бесова и др.).

Ключевые слова : минимаксное обнаружение, адаптивное обнаружение, функции большого числа переменных, критические радиусы.

Работа частично поддерживалась грантом РФФИ 08-01-00692 и грантом ведущих научных школ НШ 638.2008.1.

Для получения содержательной задачи мы также должны исключить из рассмотрения “достаточно малые” функции. Для заданных $\mathcal{F} \subset L_2^d$, $r_\varepsilon > 0$ обозначим

$$\mathcal{F}(r_\varepsilon) = \{f \in \mathcal{F} : \|f\| \geq r_\varepsilon\} \tag{1.2}$$

и будем рассматривать альтернативу $H_1 : f \in \mathcal{F}(r_\varepsilon)$. Обозначим через $\gamma_\varepsilon(\mathcal{F}, r_\varepsilon, \psi)$ сумму вероятности ошибки первого рода и максимальной вероятности ошибки второго рода теста ψ[1], то есть,

$$\gamma_\varepsilon(\mathcal{F}, r_\varepsilon, \psi) = E_{\varepsilon,0}\psi + \sup_{f \in \mathcal{F}(r_\varepsilon)} E_{\varepsilon,f}(1-\psi)$$

и пусть

$$\gamma_\varepsilon = \gamma_\varepsilon(\mathcal{F}, r_\varepsilon) = \inf_\psi \gamma_\varepsilon(\mathcal{F}, r_\varepsilon, \psi),$$

где инфимум берется по всем возможным тестам ψ. Хорошо известно (см. [5, 6]), что $\gamma_\varepsilon \in [0,1]$. Нас будет интересовать: *насколько малым может быть r_ε, чтобы $\gamma_\varepsilon \to 0$ при $\varepsilon \to 0$?*

Здесь и далее рассматриваем пределы при $\varepsilon \to 0$, если не указан другой предельный переход; соотношение $A_\varepsilon \sim B_\varepsilon$ означает, что $\lim A_\varepsilon/B_\varepsilon = 1$, а соотношение $A_\varepsilon \asymp B_\varepsilon$, что существуют такие константы $C > c > 0$, $\varepsilon_0 > 0$, что $C > A_\varepsilon/B_\varepsilon > c$ при $\varepsilon \in (0, \varepsilon_0)$.

Часто ответ может быть дан в терминах *критических радиусов*: то есть такого семейства $r_\varepsilon^* = r_\varepsilon(\mathcal{F})$, что $\gamma_\varepsilon \to 1$, если $r_\varepsilon/r_\varepsilon^* \to 0$ и $\gamma_\varepsilon \to 0$, если $r_\varepsilon/r_\varepsilon^* \to \infty$.

Задачи проверки гипотез достаточно хорошо изучены для широкого класса функциональных множеств (см. [3, 4, 1, 5, 6]). Пусть $d = 1$ и $\mathcal{F} = S_{\sigma,1}$ есть соболевский шар. Тогда асимптотика критических радиусов имеет вид, см. [3],

$$r_\varepsilon^* = r_\varepsilon(S_{\sigma,1}) \asymp \varepsilon^{4\sigma/(4\sigma+1)}. \tag{1.3}$$

Для шаров $\mathcal{F} = S_{\sigma,d}$, состоящих из σ-гладких функций d переменных, соотношение (1.3), определяющее асимптотику критических радиусов, сохраняется с заменой σ на σ/d, см. [4].

Другие ограничения регулярности соответствуют шарам $S_{\sigma,d}^{\text{ten}}$ в тензорном произведении соболевских пространств. Критические радиусы в этом случае убывают быстрее, чем для соболевских шаров:

$$r_\varepsilon(S_{\sigma,d}^{\text{ten}}) \asymp \left(\varepsilon^4 (\log(\varepsilon^{-1}))^{d-1}\right)^{\sigma/(4\sigma+1)},$$

[1] Тест есть измеримая функция наблюдений со значениями в $[0,1]$.

что совпадает с (1.3), с точностью до логарифма, см. [10]. Аналогичные результаты для задачи оценивания получены в [2, 12].

Эти результаты справедливы при фиксированных d. Если $d = d_\varepsilon \to \infty$ не слишком быстро, то асимптотика критических радиусов для соболевских шаров несколько изменяется и зависит от вида соболевской нормы. При этом логарифмическая асимптотика сохранятся. Именно, если $d = o(\log(\varepsilon^{-1}))$, то для соболевских шаров имеем

$$\log(r_\varepsilon(S_{\sigma,d})) = \frac{4\sigma\log(\varepsilon)}{4\sigma + d} + O(\log(d)),$$

а для шаров из тензорных произведений

$$\log(r_\varepsilon(S^{\mathrm{ten}}_{\sigma,d})) \sim \frac{4\sigma\log(\varepsilon)}{4\sigma + 1}.$$

Логарифмическая асимптотика имеет другой вид, если одновременно выполнены соотношения $d \gg \log(\varepsilon^{-1})$ и $\log(d) = o(\log(\varepsilon^{-1}))$, а если $\log(d) > b\log(\varepsilon^{-1})$ для некоторого $b > 0$, то критические радиусы не стремятся к нулю (см. [8, 10]).

Эти результаты иллюстрируют "проклятие размерности" в задаче проверки гипотез – порядки критических радиусов становятся весьма плохими при возрастании количества переменных.

Для преодоления "проклятия размерности" в [14] был предложен новый вариант ограничений регулярности, который оказался весьма эффективным для задач численного интегрирования, см. [11, 16], а также для задач оценивания и проверки гипотез, см. [9].

Эти ограничения соответствуют шару $\mathcal{F} = SW_{\sigma,s}$ в пространстве взвешенных тензорных произведений, которое мы будем называть пространством Слоана–Вожняковского (Sloan-Woźniakowski space). Здесь $\sigma > 0$ – параметр, характеризующий гладкость, а параметр $s > 0$ определяет последовательность "весов", которая характеризует "значимость" переменных.

Напомним определение пространств Слоана–Вожняковского с периодическими ограничениями [9]. Рассмотрим вероятностное пространство с множеством элементарных исходов – $[0,1]^\infty$, цилиндрической σ-алгеброй и вероятностной мерой – произведением мер Лебега. Пусть $L_2^\infty = L_2([0,1]^\infty)$, тогда стандартный базис Фурье в $L_2(0,1)$

$$\phi_0(t) = 1, \quad \phi_k(t) = \sqrt{2}\cos(2\pi k t),$$
$$\phi_{-k}(t) = \sqrt{2}\sin(2\pi k t) \quad \text{при} \quad k > 0$$

порождает базис Фурье $\{\phi_l(t)\}$ в L_2^∞,

$$\phi_l(t) = \prod_{k=1}^{\infty} \phi_{l_k}(t_k), \quad l = (l_1, \dots, l_k, \dots) \in \mathbb{Z}_*^\infty, \qquad t = (t_1, \dots, t_k, \dots) \in [0,1]^\infty. \tag{1.4}$$

Здесь $\mathbb{Z}_*^\infty$ состоит из таких целочисленных последовательностей $l = (l_1, \dots, l_k, \dots)$, $l_k \in \mathbb{Z}$, что $\#(k \in \mathbb{N} : l_k \neq 0) < \infty$. Рассмотрим разложение $f \in L_2^\infty$ в ряд Фурье

$$f(t) = \sum_{l \in \mathbb{Z}_*^\infty} \theta_l \phi_l(t), \quad \theta_l = (f, \phi_l).$$

Для $\sigma > 0, s > 0$, положим

$$\|f\|_{\sigma,s}^2 = \sum_{l \in \mathbb{Z}_*^\infty} \theta_l^2 c_l^2(\sigma, s), \tag{1.5}$$

где

$$c_{0,\dots,0,\dots}(\sigma, s) = 1, \quad c_l(\sigma, s) = \prod_{j \in \mathbb{N}:\, l_j \neq 0} j^s |2\pi l_j|^\sigma, \quad l \in \mathbb{Z}_*^\infty. \tag{1.6}$$

Пространство Слоана–Вожняковского $L_{\sigma,s}$ состоит из таких функций $f \in L_2^\infty$, что $\|f\|_{\sigma,s} < \infty$. Оно является гильбертовым пространством со структурой тензорного произведения

$$L_{\sigma,s} = \bigotimes_{j=1}^{\infty} L_{\sigma,j,s},$$

где гильбертово пространство $L_{\sigma,j,s}$ состоит из функций $f(t) = \sum_{k \in \mathbb{Z}} \theta_k \phi_k(t)$, $t \in [0,1]$, с нормой

$$\|f\|_{\sigma,j,s}^2 = \sum_{k \in \mathbb{Z}} b_{j,k}^2 \theta_k^2, \quad b_{j,0} = 1, \quad b_{j,k}^2 = j^{2s} |2\pi k|^{2\sigma}, \quad k \neq 0.$$

Множество $SW_{\sigma,s}$ есть эллипсоид

$$\mathcal{F} = \left\{ f(t) = \sum_{l \in \mathcal{L}} \theta_l \phi_l(t) : \sum_{l \in \mathcal{L}} c_l^2 \theta_l^2 \leqslant 1 \right\}, \tag{1.7}$$

с коэффициентами вида (1.6) и множеством индексов $\mathcal{L} = \mathbb{Z}_*^\infty$, соответствующим базису (1.4).

Точная асимптотика в задачах оценивания и проверки гипотез для таких множеств изучалась в [9]. Точные результаты [9] достаточно сложны, однако из них вытекает простая логарифмическая асимптотика критических радиусов в задаче проверки гипотез

$$\log(r_\varepsilon(SW_{\sigma,s})) \sim \frac{4\sigma^* \log \varepsilon}{4\sigma^* + 1}, \quad \sigma^* = \min(\sigma, s). \tag{1.8}$$

Эта асимптотика совпадает с логарифмической асимптотикой, получаемой из (1.3), в задаче обнаружения функции одного переменного с параметром гладкости σ^*.

В [9] также были построены асимптотически минимаксные тесты. Эти тесты зависят от параметров (σ, s), которые обычно неизвестны статистику. Возникает задача: построить тесты, не зависящие от этих параметров и обеспечивающие наилучшие критические радиусы для достаточно широкой зоны неизвестных (σ, s).

Точнее, пусть задано семейство положительных функций $r_\varepsilon = r_\varepsilon(s, \sigma)$, $(\sigma, s) \in \mathbb{R}_+^2$. Фиксируем компактное множество $K = \{(\sigma, s) \subset \mathbb{R}_+^2\}$ и рассмотрим множество альтернатив следующего вида

$$\mathcal{F}_K(r_\varepsilon) = SW_k(r_\varepsilon) = \bigcup_{(\sigma,s)\in K} \{f \in SW_{\sigma,s} : \|f\| \geqslant r_\varepsilon(s, \sigma)\}.$$

Как и раньше, положим

$$\gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon, \psi) = E_{\varepsilon,0}\psi + \sup_{f \in \mathcal{F}_K(r_\varepsilon)} E_{\varepsilon,f}(1 - \psi),$$

и

$$\gamma_\varepsilon = \gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon) = \inf_\psi \gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon, \psi).$$

Нас интересует, каким должно быть семейство $r_\varepsilon(s, \sigma)$, чтобы $\gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon) \to 1$ или $\gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon) \to 0$, и если $\gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon) \to 0$, то нас интересует структура теста ψ_ε, для которого $\gamma_\varepsilon(\mathcal{F}_K, r_\varepsilon, \psi_\varepsilon) \to 0$.

Описанную задачу называют *задачей адаптации* в проверке гипотез. Впервые она изучалась Спокойным в [15] в задаче обнаружения функции одного переменного σ с неизвестным параметром гладкости $\sigma > 0$ на основе разложения по вейвлет-базису $\{\phi_{ij},\ i = 1, \dots, 2^j,$

$j \geqslant J_0\}$. В работе [15] предлагается комбинировать набор тестов, основанных на статистиках типа χ_j^2, которые строятся по набору эмпирических коэффициентов Фурье $X_{ij} = X(\phi_{ij})$, $i = 1, \ldots, 2^j$, для вейвлет-уровней $J_0 \leqslant j \leqslant J_\varepsilon$ при достаточно большом J_ε. Показано, что эта процедура оптимальна по порядку и ведет к сравнительно малым потерям в асимптотике критических радиусов (потери соответствуют появлению дополнительного логарифмического множителя, несущественного для логарифмической асимптотики).

В случае периодических функций одной переменной эта процедура соответствует комбинации набора тестов, основанных на статистиках типа $\chi_{N_m}^2$, которые строятся по набору эмпирических коэффициентов Фурье $X_l = X(\phi_l)$, $l \in \mathcal{N}_m \cap \mathbb{Z}$, где $\mathcal{N}_m$ – симметричные интервалы вида $\mathcal{N}_m = [-N_m, N_m]$, $N_m = 2^m$, $1 \leqslant m \leqslant M_\varepsilon$, и M_ε достаточно велико. Однако для случая функций многих переменных при построении подобных процедур возникают трудности, связанные с тем, что не ясно, какими множествами мультииндексов $\mathcal{N}_m \subset \mathbb{Z}^d$ следует заменить симметричные интервалы. Эти трудности усиливаются с ростом d.

Мы преследуем в этой работе две цели:

- привести простое доказательство вида логарифмической асимптотики критических радиусов (1.8);
- указать сравнительно простые тесты для пространства Слоана–Вожняковского, не зависящие от параметров (σ, s) и обеспечивающие логарифмическую асимптотику критических радиусов (1.8).

2. Основные результаты

Следующая теорема дает нижние границы для критических радиусов.

Теорема 2.1. *Пусть*

$$r_\varepsilon = o\left(\varepsilon^{\frac{4\sigma^*}{4\sigma^*+1}}\right), \tag{2.1}$$

где $\sigma^* = \min(\sigma, s)$. Тогда $\gamma_\varepsilon(SW_{\sigma,s}, r_\varepsilon) \to 1$. Более того, последнее соотношение выполнено равномерно по любому компакту $K = \{(\sigma, s)\}$ в $\mathbb{R}_+^2$, если

$$\lim \sup_{(\sigma,s)\in K} r_\varepsilon(\sigma, s)\varepsilon^{-\frac{4\sigma^*}{4\sigma^*+1}} = 0.$$

Доказательство теоремы 2.1 приведено в разделе 3.

Опишем тесты, на которых достигается оптимальный логарифмический порядок критических радиусов. Выберем семейство

$$T_m = 2^m, \quad m = 1, \dots, M_\varepsilon, \quad M_\varepsilon \asymp \log(\varepsilon^{-1}) \log\log(\varepsilon^{-1}).$$

Пусть коэффициенты $c_l(1,1)$ определены соотношением (1.6). Положим

$$\mathcal{N}_{1,1}(T) = \{l \in \mathbb{Z}_*^\infty : c_l(1,1) \leqslant T\}, \quad N_m = \#(\mathcal{N}_{1,1}(T_m)), \tag{2.2}$$

отметим, что $N_m \geqslant N_1 = 1$. Введем статистики

$$t_{m,\varepsilon} = \frac{1}{\sqrt{2N_m}} \sum_{l \in \mathcal{N}_{1,1}(T_m)} ((X(\phi_l)/\varepsilon)^2 - 1), \quad t_\varepsilon = \max_{1 \leqslant m \leqslant M_\varepsilon} t_{m,\varepsilon},$$

Фиксируем $B > 2$ и рассмотримм семейство тестов

$$\psi_\varepsilon = \mathbb{1}_{\{t_\varepsilon > L_\varepsilon\}}, \quad L_\varepsilon = \sqrt{B \log M_\varepsilon}.$$

Пусть $K = \{(\sigma, s)\}$ есть компактное множество в $\mathbb{R}_+^2$ и $r_\varepsilon(\sigma, s)$ – функция на $\mathbb{R}_+^2$. Из теоремы 2.1 следует, что, соотношение $\liminf \gamma_\varepsilon(SW_K, r_\varepsilon, \psi_\varepsilon) < 1$ может быть выполнено лишь в случае, если

$$\lim \inf_{(\sigma,s) \in K} r_\varepsilon(\sigma, s) \varepsilon^{-\frac{4\sigma^*}{4\sigma^*+1}} > 0.$$

Теорема 2.2. *Пусть существует такое $\delta > 0$, что $\forall\ (\sigma, s) \in K$*

$$r_\varepsilon(\sigma, s) \varepsilon^{-\frac{4\sigma^*}{4\sigma^*+1}} \geqslant \varepsilon^{-\delta}, \tag{2.3}$$

где $\sigma^* = \min(\sigma, s)$. Тогда

$$\gamma_\varepsilon(SW_K, r_\varepsilon, \psi_\varepsilon) \to 0.$$

Доказательство теоремы 2.2 приведено в разделе 4.

Обсудим структуру множеств

$$\mathcal{N}_T = \mathcal{N}_{1,1}(T) = \Big\{ l \in \mathbb{Z}_*^\infty : \prod_{j \in \mathbb{N} :\, l_j \neq 0} (2\pi |l_j| j) \leqslant T \Big\}$$

(произведение по пустому множеству полагаем равным единице).

Заметим, что $\mathbb{Z}_*^\infty = \bigcup_u \mathbb{Z}_u$, где объединение берется по всем конечным подмножествам $u = \{j_1, \dots, j_m\} \subset \mathbb{N}$ и $\mathbb{Z}_u$ состоит из таких целочисленных последовательностей l, что $l_j \neq 0$, если и только если $j \in u$ (множество $\mathbb{Z}_u$ изоморфно $(\mathbb{Z} \setminus \{0\})^m$, $0 \leqslant m = \#(u) < \infty$, при $m = 0$ подмножество $u = \emptyset$ и $\mathbb{Z}_u = (0, \dots, 0, \dots)$). Это влечет $\mathcal{N}_T = \bigcup_u \mathcal{N}_{u,T}$, где $\mathcal{N}_{u,T}$ являются множествами "гиперболического" типа:

$$\mathcal{N}_{u,T} = \Big\{ l \in \mathbb{Z}_u : \prod_{j\in u} |l_j| \leqslant T/\gamma(u) \Big\}, \quad \text{где} \quad \gamma(u) = \prod_{j\in u} (2\pi j).$$

3. Доказательство теоремы 2.1

Переходя к последовательностям, достаточно рассмотреть случаи последовательностей $\varepsilon_n \to 0$ и таких $(\sigma_{\varepsilon_n}, s_{\varepsilon_n})$, что $\sigma_{\varepsilon_n} \to \sigma^0 > 0$, $s_{\varepsilon_n} \to s^0 > 0$ и $\sigma = \sigma_{\varepsilon_n} \leqslant s_{\varepsilon_n} = s$ или $\sigma = \sigma_{\varepsilon_n} \geqslant s_{\varepsilon_n} = s$. Соотношение (2.1) выполнено для этих последовательностей.

Пусть $\sigma \leqslant s$. Рассмотрим множество W_σ^1, состоящее из функций одного переменного

$$f(t) = \sum_{k=-\infty}^{\infty} \theta_k \phi_k(t) \quad \text{таких, что} \quad \theta_0^2 + \sum_{k\in\mathbb{Z},\ k\neq 0} (2\pi|k|)^{2\sigma} \theta_k^2 \leqslant 1.$$

При любом $s > 0$ справедливо вложение $W_\sigma^1 \subset SW_{\sigma,s}$, которое влечет неравенство $\gamma_\varepsilon(W_\sigma^1, r_\varepsilon) \leqslant \gamma_\varepsilon(SW_{\sigma,s}, r_\varepsilon)$. Множество W_σ^1 есть эллипсоид с коэффициентами $c_k \asymp |k|^\sigma$ (то есть такими, что $b \leqslant c_k |k|^{-\sigma} \leqslant B$ для всех $k \in \mathbb{Z} \setminus \{0\}$ при некоторых постоянных $0 < b < B < \infty$). Для эллипсоидов с такой асимптотикой коэффициентов хорошо известно (см. [5, 6, раздел 3.2.1]; рассуждения [6] переносятся на случай последовательностей $\sigma_{\varepsilon_n} \to \sigma^0 > 0$), что $\gamma_\varepsilon(W_\sigma^1, r_\varepsilon) \to 1$ при выполнении (2.1). Таким образом, теорема 2.1 справедлива при $\sigma \leqslant s$.

Пусть $\sigma \geqslant s$. Рассмотрим множество W_s состоящее из функций

$$f(t_1, \dots, t_j, \dots) = \sum_{j=1}^{\infty} \theta_j \phi_1(t_j)$$

$$\text{таких, что} \quad (2\pi)^{2\sigma} \sum_{j=1}^{\infty} j^{2s} \theta_j^2 \leqslant 1. \quad (3.1)$$

При любом $\sigma > 0$ справедливо вложение $W_s \subset SW_{\sigma,s}$, откуда $\gamma_\varepsilon(W_s, r_\varepsilon) \leqslant \gamma_\varepsilon(SW_{\sigma,s}, r_\varepsilon)$. Множество W_s есть эллипсоид с коэффициентами $c_j \asymp j^s$, $j \in \mathbb{N}$. Аналогично предыдущему случаю отсюда следует, что $\gamma_\varepsilon(W_s, r_\varepsilon) \to 1$ при (2.1). Теорема 2.1 доказана. □

4. Доказательство теоремы 2.2

Обозначим

$$\mathcal{N}_{\sigma,s}(T) = \{l \in \mathbb{Z}_*^\infty : c_l(\sigma, s) \leqslant T\}, \quad N_{\sigma,s}(T) = \#(\mathcal{N}_{\sigma,s}(T)).$$

Лемма 4.1. *Равномерно по любому компакту $K = \{(\sigma, s)\} \subset \mathbb{R}_+^2$,*

$$\log\big(N_{\sigma,s}(T)\big) \sim \log(T)/\sigma^*, \quad \textit{если} \quad T \to \infty, \quad \sigma^* = \min(\sigma, s).$$

Доказательство леммы 4.1. Покажем сначала, что

$$\log\big(N_{\sigma,s}(T)\big) \geqslant \log(T)/\sigma^* + O(1). \tag{4.1}$$

Доказательство (4.1) аналогично доказательству теоремы 2.1. Пусть $\sigma \leqslant s$. Множество $\mathcal{N}_{\sigma,s}(T)$ содержит точки $l^{(k)} = (k, 0, \dots, 0, \dots)$, $k \in \mathbb{Z}$, $k \neq 0$, $(2\pi|k|)^\sigma \leqslant T$. Это влечет (4.1). Пусть $s > \sigma$. Множество $\mathcal{N}_{\sigma,s}(T)$ содержит точки $l^{(j)} = (0, \dots, 0, 1, 0, \dots)$ (единица стоит на j-ом месте), $j \in \mathbb{N}$, $(2\pi)^\sigma j^s \leqslant T$. Это также влечет (4.1).

Покажем теперь, что

$$\log\big(N_{\sigma,s}(T)\big) \leqslant \log(T)/\sigma^*(1 + o(1)). \tag{4.2}$$

Доказательство аналогично доказательству предложения 6.1 в [10]. Представим коэффициенты $c_l(\sigma, s)$ в (1.6) в виде

$$c_l(\sigma, s) = \exp(S(l)),$$

где

$$\begin{aligned} S(l) &= \sum_{j=1}^\infty Y_j(l_j), \\ Y_j(k) &= \begin{cases} 0, & k = 0, \\ s\log j + \sigma\log(2\pi|k|), & k \neq 0, \quad k \in \mathbb{Z}. \end{cases} \end{aligned} \tag{4.3}$$

Аналогично [9], для $h > 1/\sigma^*$, $j \in \mathbb{N}$ определим вероятностную меру $P_{h,j}$ на $\mathbb{Z}$, полагая

$$P_{h,j}(k) = \exp(-hY_j(k) - Z_j(h)),$$
$$Z_j(h) = \log\left(\sum_{k\in\mathbb{Z}} \exp(-hY_j(k))\right) > 0,$$

где $Y_j(k)$ определены в (4.3). Эта мера корректно определена при $h\sigma > 1$, так как

$$1 < \sum_{k\in\mathbb{Z}} \exp(-hY_j(k)) = 1 + j^{-sh}G(h) < \infty, \quad G(h) = 2\sum_{k=1}^{\infty}(2\pi k)^{-h\sigma}.$$

Рассмотрим меру $P_h = \prod_{j\in\mathbb{N}} P_{h,j}$. Для любого $l = \{l_j\} \in \mathbb{Z}_*^\infty$ имеем

$$P_h(l) = \prod_{j\in\mathbb{N}} P_{h,j}(l_j) = \exp\Big(-h\sum_{j\in\mathbb{N}} Y_j(l_j) - Z(h)\Big) = \exp(-hS(l) - Z(h)),$$

где $Z(h) = \sum_{j=1}^{\infty} Z_j(h)$. Так как $sh > 1$, величина $Z(h)$ ограничена

$$0 < Z(h) = \sum_{j=1}^{\infty} Z_j(h) = \sum_{j=1}^{\infty} \log(1 + j^{-sh}G(h)) < G(h)\sum_{j=1}^{\infty} j^{-sh} < \infty.$$

Легко проверить, что P_h сосредоточена на $\mathbb{Z}_*^\infty$, то есть $P_h(\mathbb{Z}_*^\infty) = 1$. Полагая $H = \log(T)$ и обозначая через E_h математическое ожидание по мере P_h, получим (ср. с [10])

$$\begin{aligned} N_{\sigma,s}(T) &= \#\{l : c_l(\sigma,s) \leqslant T\} = \#\{l : S(l) \leqslant H\} \\ &= e^{Z(h)+hH} \sum_{l:S(l)\leqslant H} \exp(h(S(l) - H))P_h(l) \\ &= e^{Z(h)+hH} E_h \exp(h(S - H))\mathbb{1}_{\{S\leqslant H\}} \leqslant e^{Z(h)+hH}. \end{aligned}$$

Для любого $\delta > 0$ положим $h = 1/\sigma^* + \delta$, тогда при достаточно больших T получим

$$\frac{\log(N_{\sigma,s}(T))}{\log(T)} \leqslant 1/\sigma^* + \delta + \frac{Z(1/\sigma^* + \delta)}{\log(T)} < 1/\sigma^* + 2\delta.$$

Отсюда следует (4.2). □

4.1. Вероятности ошибок первого рода

Отметим, что

$$
\begin{aligned}
E_{\varepsilon,0}\psi_\varepsilon = P_{\varepsilon,0}(t_\varepsilon > L_\varepsilon) \leqslant \sum_{m=1}^{M_\varepsilon} P_{\varepsilon,0}(t_{\varepsilon,m} > L_\varepsilon),\\
L_\varepsilon = \sqrt{B\log(M_\varepsilon)}, \quad B > 2.
\end{aligned} \tag{4.4}
$$

Статистики $t_{m,\varepsilon}$ относительно меры $P_{\varepsilon,0}$ совпадают по распределению с величиной $\widehat{\chi}^2_N = (\chi^2_N - N)/\sqrt{2N}$, где χ^2_N имеет хи-квадрат распределение с $N = N_m$ степенями свободы, величины N_m определены в (2.2). Для любого $z > 0$ справедливо неравенство

$$
\log(P(\chi^2_N > N(1+z))) \leqslant -\frac{N}{2}(z - \log(1+z)). \tag{4.5}
$$

Неравенство (4.5) следует из неравенства Маркова и вычисления экспоненциальных моментов χ^2_N, см., например, [7, неравенство (5.22)]. Отсюда следует, что

$$
\log(P(\widehat{\chi}^2_N > L)) \leqslant -L^2/2(1+o(1)) \ \text{ при } \ L = o(\sqrt{N}), \quad L \to \infty, \tag{4.6}
$$

и можно выбрать такое $C > 0$, что для любых $N \geqslant 1$, $L > 0$, выполнено

$$
\log(P(\widehat{\chi}^2_N > L)) \leqslant -\begin{cases} CL^2, & \text{если } \ L \leqslant \sqrt{N},\\ CL\sqrt{N}, & \text{если } \ L > \sqrt{N}. \end{cases} \tag{4.7}
$$

Напомним, что $N_m \geqslant 1$. Из леммы 4.1 следует, что

$$
\log N_m \sim \log T_m = m\log(2). \tag{4.8}
$$

Представим сумму в (4.4) в виде трех сумм

$$
\sum_{m=1}^{M_\varepsilon} P_{\varepsilon,0}(t_{m,\varepsilon} > L_\varepsilon) = \sum_{\substack{L_\varepsilon < 2^{m/4},\\ 1\leqslant m\leqslant M_\varepsilon}} + \sum_{\substack{2^{m/4}\leqslant L_\varepsilon\leqslant\sqrt{N_m},\\ 1\leqslant m\leqslant M_\varepsilon}} + \sum_{\substack{\sqrt{N_m}<L_\varepsilon,\\ 1\leqslant m\leqslant M_\varepsilon}}.
$$

В силу (4.8), мы можем применить (4.6) к слагаемым в первой сумме. Получим при $M_\varepsilon \to \infty$,

$$
\sum_{L_\varepsilon < 2^{m/4},\ 1\leqslant m\leqslant M_\varepsilon} P_{\varepsilon,0}(t_{\varepsilon,m} > L_\varepsilon) \leqslant \frac{M_\varepsilon}{(M_\varepsilon)^{B/2(1+o(1))}} \to 0.
$$

Рассмотрим вторую и третью суммы. В силу (4.8), число слагаемых в этих суммах ограничено величиной $O(\log\log(M_\varepsilon))$, и мы можем использовать (4.7) для оценки каждого слагаемого этих сумм. При $M_\varepsilon \to \infty$, получим

$$\sum_{\substack{2^{m/4} \leqslant L_\varepsilon \leqslant \sqrt{N_m},\\ 1\leqslant m \leqslant M_\varepsilon}} \leqslant \sum_{2^{m/4} \leqslant L_\varepsilon \leqslant \sqrt{N_m}} P_{\varepsilon,0}(t_{\varepsilon,m} > L_\varepsilon) \leqslant \frac{O(\log\log(M_\varepsilon))}{M_\varepsilon^{BC}} \to 0,$$

$$\sum_{\substack{\sqrt{N_m} < L_\varepsilon,\\ 1\leqslant m \leqslant M_\varepsilon}} \leqslant \sum_{\sqrt{N_m} < L_\varepsilon} P_{\varepsilon,0}(t_{\varepsilon,m} > L_\varepsilon) \leqslant \frac{O(\log\log(M_\varepsilon))}{\exp(C\sqrt{B\log(M_\varepsilon)}\,)} \to 0.$$

Отсюда следует, что $E_{\varepsilon,0}\psi_\varepsilon \to 0$. □

4.2. Вероятности ошибок второго рода

Пусть выполнено (2.3). Нужно показать, что

$$\sup_{f\in SW_{\sigma,s}(r_\varepsilon(\sigma,s))} E_{\varepsilon,f}(1-\psi_\varepsilon) \to 0.$$

Заметим, что

$$E_{\varepsilon,f}(1-\psi_\varepsilon) = P_{\varepsilon,f}(t_\varepsilon \leqslant L_\varepsilon) \leqslant \min_{1\leqslant m\leqslant M_\varepsilon} P_{\varepsilon,f}(t_{m,\varepsilon} \leqslant L_\varepsilon).$$

Достаточно найти такое целое $m = m_\varepsilon(\sigma,s) \in [1, M_\varepsilon]$, что

$$\sup_{f\in SW_{\sigma,s}(r_\varepsilon(\sigma,s))} P_{\varepsilon,f}(t_{m,\varepsilon} \leqslant L_\varepsilon) \to 0.$$

Выберем такое m, что

$$T_m = 2^m \asymp \varepsilon^{-4/(4\sigma^*+1)}, \tag{4.9}$$

и отметим, что $m \asymp \log(\varepsilon^{-1}) \in [1, M_\varepsilon]$ равномерно по любому компакту $K = \{(\sigma,s)\} \subset \mathbb{R}^2_+$. Из неравенства $c_l(\sigma,s) \geqslant (c_l(1,1))^{\sigma^*}$ следует, что

$$\mathcal{N}_{\sigma,s}(T_m^{\sigma^*}) \subset \mathcal{N}_{1,1}(T_m), \quad N_{\sigma,s}(T_m^{\sigma^*}) \leqslant N_{1,1}(T_m) = N_m. \tag{4.10}$$

Пусть $f \in SW_{\sigma,s}(r_\varepsilon(\sigma,s))$, тогда

$$f(t) = \sum_{l\in\mathbb{Z}_*^\infty} \theta_l \phi_l(t): \quad \sum_{l\in\mathbb{Z}_*^\infty} c_l^2(\sigma,s)\theta_l^2 \leqslant 1, \quad \sum_{l\in\mathbb{Z}_*^\infty} \theta_l^2 \geqslant r_\varepsilon^2(\sigma,s).$$

Распределение статистик $t_{m,\varepsilon}$ относительно меры $P_{\varepsilon,f}$ были изучены, например, в [5, III] и [6, раздел 3.1.2]. Имеем

$$\begin{aligned} E_{\varepsilon,f}(t_{m,\varepsilon}) &= \frac{1}{\varepsilon^2\sqrt{2N_m}} \sum_{l\in\mathcal{N}_{1,1}(T_m)} \theta_l^2, \\ \operatorname{Var}_{\varepsilon,f}(t_{m,\varepsilon}) &= 1 + \frac{2}{\varepsilon^2 N_m} \sum_{l\in\mathcal{N}_{1,1}(T_m)} \theta_l^2 = 1 + \frac{2E_{\varepsilon,f}(t_{m,\varepsilon})}{\sqrt{N_m/2}} \\ &= 1 + o(E_{\varepsilon,f}(t_{m,\varepsilon})). \end{aligned}$$

Используя (2.3), (4.9), (4.10), получим

$$\begin{aligned} \sum_{l\in\mathcal{N}_{1,1}(T_m)} \theta_l^2 &\geqslant \sum_{l\in\mathcal{N}_{\sigma,s}(T_m^{\sigma^*})} \theta_l^2 = \sum_{l\in\mathbb{Z}_*^\infty} \theta_l^2 - \sum_{l:c_l(\sigma,s)>T_m^{\sigma^*}} \theta_l^2 \\ &\geqslant r_\varepsilon^2(\sigma,s) - T_m^{-2\sigma^*} \sum_{l:c_l(\sigma,s)>T_m^{\sigma^*}} c_l^2(\sigma,s)\theta_l^2 \\ &\geqslant r_\varepsilon^2(\sigma,s) - T_m^{-2\sigma^*} \geqslant \varepsilon^{\frac{8\sigma^*}{4\sigma^*+1}-2\delta} - O(\varepsilon^{\frac{8\sigma^*}{4\sigma^*+1}}) \sim \varepsilon^{\frac{8\sigma^*}{4\sigma^*+1}-2\delta}. \end{aligned}$$

Из леммы 4.1 следует, что

$$N_m = T_m^{1+o(1)} = \varepsilon^{-\frac{4+o(1)}{4\sigma^*+1}}.$$

Равномерно по $f \in SW_{\sigma,s}(r_\varepsilon(\sigma,s))$, имеем

$$E_{\varepsilon,f}(t_{m,\varepsilon}) \geqslant \varepsilon^{-2+\frac{2+o(1)}{4\sigma^*+1}+\frac{8\sigma^*}{4\sigma^*+1}-2\delta}/\sqrt{2} \geqslant \varepsilon^{-\delta}.$$

Используя неравенство Чебышева, получим

$$\begin{aligned} P_{\varepsilon,f}(t_{m,\varepsilon} \leqslant L_\varepsilon) &= P_{\varepsilon,f}(E_{\varepsilon,f}(t_{m,\varepsilon}) - t_{m,\varepsilon} \geqslant E_{\varepsilon,f}(t_{m,\varepsilon}) - L_\varepsilon) \\ &\leqslant \frac{\operatorname{Var}_{\varepsilon,f}(t_{m,\varepsilon})}{(E_{\varepsilon,f}(t_{m,\varepsilon}) - L_\varepsilon)^2} = o(\varepsilon^\delta). \end{aligned}$$

Все оценки, полученные выше, равномерны по любому компакту $K = \{(\sigma,s)\} \subset \mathbb{R}_+^2$. Теорема 2.2 доказана. □

Литература

1. М. С. Ермаков, *Минимаксное обнаружение сигнала в гауссовском белом шуме.* — Теория вероятн. ее примен. **35**, 4 (1990), 704–715.
2. I. A. Ibragimov, R. Z. Khasminskii, *Some estimation problems on Infinite dimensional Gaussian white noise.* — In: Festschrift for Lusien Le Cam. Research papers in Probability and Statictics. Springer-Verlag, NY (1997), pp. 275–296.
3. Ю. И. Ингстер, *О минимаксном непараметрическом обнаружении сигнала в гауссовском белом шуме.* — Проблемы передачи информации **18**, 2 (1982), 61–73.
4. Yu. I. Ingster, *Asymptotically minimax testing of nonparametric hypotheses.* — In: Proc. 4th Vilnius Conference on Probab. Theory and Math. Stat., VNU Science Press, **1** (1987), pp. 553–573.
5. Yu. I. Ingster, *Asymptotically minimax hypothesis testing for nonparametric alternatives.* I, II, III. — Math. Methods Statist. **2** (1993), 85–114, 171–189, 249–268.
6. Yu. I. Ingster, I. A. Suslina, *Nonparametric Goodness-of-Fit Testing under Gaussian Model.* — Lect Notes Statist. Vol. **169**, Springer-Verlag, New York, 2002.
7. Yu. I. Ingster, I. A. Suslina, *Nonparametric hypothesis testing for small type* I *errors.* I . — Math. Methods Statist. **13** (2004), 409–459.
8. Yu. I. Ingster, I. A. Suslina, *On estimation and detection of smooth function of many variables.* — Math. Methods Statist. **14** (2005), 299–331.
9. Yu. I. Ingster, I. A. Suslina, *Estimation and detection of high-variable function from Sloan–Woźniakowski space.* — Math. Methods Statist. **16** (2007), 318–353.
10. Ю. И. Ингстер, И. А. Суслина, *Оценивание и проверка гипотез для функций из тензорных произведений пространств.* — Зап. научн. семин. ПОМИ **351** (2007), 180–218.
11. F. Y. Kuo, J. H. Sloan, *Lifting the curse of dimensionality.* — Notices Amer. Math. Soc. **52** (2005), 1320–1329.
12. Y. Lin, *Tensor product space ANOVA model.* — Ann. Statist. **28** (2000), 734–755.
13. А. В. Скороход, *Интегрирование в гильбертовом пространстве.* Наука, М., 1975.
14. I. H. Sloan, H. Woźniakowski, *When are quazi-Monte Carlo algorithms efficient for high dimensional integrals?* — J. Complexity **14** (1998), 1–33.
15. V. G. Spokoiny, *Adaptive hypothesis testing using wavelets.* — Ann. Statist. **24** (1996), 2477–2498.
16. H. Woźniakowski, *Tractability of multivariate problems for weighted spaces of functions.* — In: Appoximation and Probability, Banach Center Publications **72**, Institute of Mathematics, Polish Academy of Science, Warszawa (2006).

Ingster Yu. I., Suslina I. A. Adaptive detection of functions of large number of variables.

A major difficulty arising in statistics of multi-variable functions is “the curse of dimensionality”: the rates of accuracy in estimation and separation rates in detection problems behave poorly when the number of variables increases. This difficulty arises for most popular functional classes such as Sobolev or Hölder balls.

In the paper [9], it was considered functional classes of a new type, first introduced by Sloan and Woźniakowski in [14]. These classes are the balls $\mathcal{F}_{\sigma,s}$ in the weighed tensor product spaces that are characterized by two parameters: $\sigma > 0$ is a "smoothness" parameter, and $s > 0$ determines the weight sequence which characterizes "importance" of the variables. In particular, it was shown in [9] that under the white Gaussian noise model, the log-asymptotics of separation rates in detection are similar to those for one-variable functions of the smoothness $\sigma^* = \min(s, \sigma)$ independently of the original problem dimensions; thus the curse of dimensionality is "lifted." However the test procedure depends on parameters (σ, s) which are unknown typically.

In this paper, we propose a common test procedure that does not depend on parameters (σ, s) and provides the same log-asymptotics of separation rates uniformly over any compact set of parameters (σ, s). Also we give independent simple proof of the log-asymptotics of separation rates in the problem.

Поступило 10 октября 2009 г.

С. Петербургский государственный
электротехнический университет (ЛЭТИ),
ул. проф. Попова 5,
197376 Санкт-Петербург, Россия

E-mail: yurii_ingster@mail.ru

С.Петербургский государственный
университет информационных технологий,
механики и оптики,
Кронверкский пр. 49,
197101 Санкт-Петербург, Россия

E-mail: isuslina@mail.ru