Записки научных
семинаров ПОМИ
Том 339, 2006 г.

С. В. Решетов

# НИЖНИЕ ГРАНИЦЫ ДЛЯ РИСКА В ЗАДАЧЕ ОЦЕНИВАНИЯ ПСЕВДО–ПЕРИОДИЧЕСКОЙ ФУНКЦИИ, НАБЛЮДАЕМОЙ НА ФОНЕ БЕЛОГО ШУМА

## 1. Введение

Пусть на растущем отрезке $[-T, T]$ наблюдается процесс $Y(t)$,

$$dY(t) = s(t)dt + \sigma dW(t). \tag{1}$$

Здесь $s \in \mathcal{L}_*$ – неизвестная, подлежащая оцениванию функция, $W(t)$ – винеровский процесс, $\mathcal{L}_*$ – компактное подмножество Банахова пространства $\mathcal{L}(\Lambda)$ псевдо-периодических функций со спектральным множеством $\Lambda$. Последнее означает, что

$$s(t) = \sum_{u \in \Lambda} a(u) e^{\mathrm{iut}}, \quad \sum_{u \in \Lambda} |a(u)|^2 < \infty,$$

соответствующая Банахова норма задается соотношением

$$\|s\|_{\mathcal{L}}^2 = \sup_x \int_x^{x+1} |s(t)|^2 \, dt.$$

Предполагается, что $\Lambda$ – счетнное множество, удовлетворяющее условию

$$\tau = \inf_{u \neq v;\, u,v \in \Lambda} |u - v| > 0.$$

Параметрическое множество $\mathcal{L}_*$ состоит из функций $s \in \mathcal{L}(\Lambda)$, выделяемых условием

$$\sum_{u \in \Lambda} |a(u)|^2 |u|^{2\beta} \leq C. \tag{2}$$

Задача состоит в построении нижней границы для риска

$$R_T(\widehat{s}_T) = \sup_{s \in \mathcal{L}_*} \mathbf{E} \|s - \widehat{s}_T\|_{\mathcal{L}}.$$

при использовании произвольной оценки $\widehat{s}_T$, построенной по наблюдениям $Y(t), t \in [-T, T]$. Основной результат работы состоит в следующей теореме.

**Теорема 1.** *Пусть выполнено условие* (2). *Тогда существует такая положительная константа $L$, что при использовании произвольной оценки $\widehat{s}_T$, построенной по наблюдениям $Y(t), t \in [-T, T]$*

$$R_T(\widehat{s}_T) \geq \left(\frac{1}{T}\right)^{\frac{\beta}{2\beta+1}}. \tag{3}$$

Обычным орудием для построения таких границ является неравенство Фано.

## 2. Неравенство Фано

Пусть $\theta^*, X$ случайные элементы, принимающие значения в конечном множестве $\Theta = \{\theta_1, \ldots, \theta_N\}$ и измеримом пространстве $(\mathcal{X}, \mathcal{B})$ соответственно, $\widehat{\theta} = \psi(X)$ оценка $\theta^*$ по наблюдению $X$ со значениями в $\Theta$. Имеет место неравенство, называемое неравенством Фано,

$$P\{\theta^* \neq \widehat{\theta}\} \geq \frac{H(\theta^*)}{\ln(N-1)} - \frac{I(\theta^*, X) + \ln 2}{\ln(N-1)}. \tag{4}$$

Здесь $H(\theta^*)$ – энтропия разбиения, порождённого случайным элементом $\theta^*$:

$$H(\theta^*) = -\sum_{j=1}^{N} P\{\theta^* = \theta_j\} \ln P\{\theta^* = \theta_j\}.$$

Величина $I(\theta^*, X)$ - количество информации, содержащееся в $X$ относительно $\theta^*$:

$$I(\theta^*, X) = \mathbf{E} \ln \frac{d\, P_{\theta^*, X}}{d\,(P_{\theta^*} \times P_X)},$$

где $P_{\theta^*, X}$ – мера, порождённая вектором $(\theta^*, X)$, $P_{\theta^*}, P_X$ – распределения $\theta^*$ и $X$, а $P_{\theta^*} \times P_X$ – продакт-мера. Величина $I(\theta^*, X) = \infty$, если меры $P_{\theta^*, X}$ и $P_{\theta^*} \times P_X$ не являются взаимно абсолютно непрерывными.

Теперь предположим, что наблюдается случайный элемент $X$ с неизвестным распределением $P_\theta, \theta \in \Theta = \{\theta_1, \ldots, \theta_N\}$. Для простоты изложения предположим, что распределения $P_\theta$ имеют плотности $f_\theta$ по отношению к некоторой мере $\mu$. Распределение случайного вектора $(\theta^*, X)$ зададим соотношением

$$P\{X \in A, \theta^* = \theta_j\} = \frac{1}{N} \int\limits_A f_{\theta_j}(x)\, d\mu. \tag{5}$$

Мы хотим оценить величину $\theta$ по наблюдению $X$. В качестве функции потерь возьмем величину

$$l(\widehat{\theta}, \theta) = \begin{cases} 1, & \text{если} \quad \widehat{\theta} \neq \theta \\ 0, & \text{иначе.} \end{cases}$$

Тогда риск $R_l(\widehat{\theta})$ от использования оценки $\widehat{\theta}$ может быть оценен снизу следующим образом

$$R_l(\widehat{\theta}) = \sup_{\theta \in \Theta} P_\theta \left\{\widehat{\theta} \neq \theta\right\} \geq P\left\{\widehat{\theta} \neq \theta^*\right\}. \tag{6}$$

Стало быть, неравенство Фано дает нам при таком способе измерения потерь оценку

$$R_l(\widehat{\theta}) \geq \frac{H(\theta^*)}{\ln(N-1)} - \frac{I(\theta^*, X) + \ln 2}{\ln(N-1)}. \tag{7}$$

Проведем некоторые упрощения. Во-первых,

$$H(\theta^*) \leq \frac{1}{N}. \tag{8}$$

Пусть вероятностные меры $P, Q$ абсолютно непрерывны относительно $\sigma-$конечной меры $\mu$,

$$f = \frac{dP}{d\mu}, q = \frac{dQ}{d\mu}.$$

Величина

$$K(P,Q) = K(f,q) = \begin{cases} \int f \ln \frac{f}{q}\, d\mu, & \text{если меры } P \text{ и } Q \quad \text{взаимно} \\ & \text{абсолютно непрерывны} \\ \infty, & \text{иначе} \end{cases}$$

называется расстоянием Кульбака–Лейблера между $P$ и $Q$.

Исходя из (5) и определения величины $I(\theta^*, X)$ получаем

$$I(\theta^*, X) = \frac{1}{N}\sum_{j=1}^{N}\int f_{\theta_j} \ln \frac{f_{\theta_j}}{\frac{1}{N}\sum\limits_{i=1}^{N} f_{\theta_i}}\, d\mu = \frac{1}{N}\sum_{j=1}^{N} K(f_{\theta_j}, \overline{f}),$$

где $\overline{f} = \frac{1}{N}\sum\limits_{i=1}^{N} f_{\theta_i}$. Далее,

$$\frac{1}{N}\sum_{j=1}^{N} K(f_{\theta_j}, \overline{f}) = \frac{1}{N}\sum_{j=1}^{N} K(f_{\theta_j}, f_{\theta_k}) - K(\overline{f}, f_{\theta_k}) \le$$
$$\le \sup_{j,k} K(f_{\theta_j}, f_{\theta_k}) = \sup_{j,k} K(P_{\theta_j}, P_{\theta_k}).$$

Отсюда и из (8) получаем

$$R_l(\widehat{\theta}) \ge 1 - \frac{\sup\limits_{j,k} K(P_{\theta_j}, P_{\theta_k}) + \ln 2}{\ln(N-1)}. \tag{9}$$

Теперь мы готовы сформулировать нужный нам результат, касающийся оценки снизу для функции риска. Пусть $X$ — случайный элемент с распределением $P_\theta$, $\theta \in \Theta$, $\widehat{\theta} = \widehat{\theta}(X)$ – оценка $\theta$ со значениями из $\Theta$. Предположим, что на $\Theta$ задана псевдо-метрика $d$, функция потерь $L(\widehat{\theta}, \theta) = d(\widehat{\theta}, \theta)$. Далее, пусть $\Theta_\varepsilon = \{\theta_1, \ldots, \theta_N\}$ конечное подмножество множества $\Theta$, состоящее из $\varepsilon-$различимых точек:

$$d(\theta_i, \theta_j) > \varepsilon, \quad i \ne j.$$

**Лемма 1.** *Имеет место оценка*

$$R(\widehat{\theta}) = \sup_{\theta\in\Theta} \mathbf{E}_\theta\, d(\widehat{\theta}, \theta) \ge \frac{\varepsilon}{2}\left(1 - \frac{\sup\limits_{j,k} K(P_{\theta_j}, P_{\theta_k}) + \ln 2}{\ln(N-1)}\right). \tag{10}$$

**Доказательство.** Пусть $\tilde{\theta}$ – точка из множества $\Theta_\varepsilon = \{\theta_1, \ldots, \theta_N\}$, ближайшая к $\widehat{\theta}$. Очевидно,

$$d(\widehat{\theta}, \theta_j) \ge \frac{\varepsilon}{2},$$

если $\theta_j \neq \tilde{\theta}$. А потому

$$R(\widehat{\theta}) \geq \sup_{\theta\in\Theta_\varepsilon} \mathbf{E}_\theta\, d(\widehat{\theta}, \theta) \geq \frac{\varepsilon}{2} \sup_{\theta\in\Theta_\varepsilon} R_l(\tilde{\theta}). \tag{11}$$

Отсюда и из (9) получаем (11).

## 3. Пространство псевдо-периодических функций

Нас интересует банахово пространство функций $s(t)$, заданных на всей вещественной прямой, со следующей банаховой нормой:

$$\|s(t)\|^2_{\mathcal{L}} = \sup_{|x|\leq 1} \int_x^{x+1} |s(t)|^2 dt$$

Пусть $\Lambda$ – не более чем счётное множество на прямой, такое что

$$\sup_{u,v\in\Lambda, u\neq v} |u - v| \geq \tau > 0. \tag{12}$$

Обозначим через $\mathcal{L}(\Lambda)$ множество функций, представимых в виде

$$s(t) = \sum_{u\in\Lambda} a(u) e^{iut}, \quad \text{где} \quad \sum_{u\in\Lambda} |a(u)|^2 < \infty.$$

В прстранстве $\mathcal{L}(\Lambda)$ введём гильбертову норму

$$\|s(t)\|^2_L = \sum_{u\in\Lambda} |a(u)|^2.$$

По теореме Винера–Пэли (Н. Винер, Р.Пэли (1964)) при условии (12) существует такое $A$, зависящее от $\tau$, что при $T > A$ и некотором $C > 0$ выполняется следующее:

$$\inf_y \int_y^{y+1} |s(t)|^2 dt > C \sup_y \int_y^{y+1} |s(t)|^2 dt.$$

Отсюда, существуют такие константы $K_1 > K_2 > 0$, что при достаточно большом $T$

$$K_2 \|s(t)\|_{\mathcal{L}} \leq \frac{1}{2T} \int_{-T}^{T} |s(t)|^2\, dt \leq K_1 \|s(t)\|_{\mathcal{L}}, \quad s \in \mathcal{L}(\Lambda). \tag{13}$$

Отсюда же следует, что существуют такие константы $C_1 > C_2 > 0$, что

$$C_2\|s(t)\|_L \leq \|s(t)\|_{\mathcal{L}} \leq C_1\|s(t)\|_L, \; s \in \mathcal{L}(\Lambda). \tag{14}$$

Нас будут интересовать компактные подмножества

$$\mathcal{L}_* = \mathcal{L}_*(\beta, C_{\mathcal{L}_*}) \subset \mathcal{L}(\Lambda),$$

выделяемые условием:

$$\sum_{u\in\Lambda} |a(u)|^2|u|^{2\beta} < C^2_{\mathcal{L}_*}, \text{ где } \quad \beta = r + \alpha, \quad \text{а} \quad r\text{– целая часть } \beta.$$

Будем говорить, что точки множества $B$ $\varepsilon$-различимы (в смысле метрики пространства $\mathcal{L}$), если

$$\inf_{x,y\in B, x\neq y} \|x-y\|_{\mathcal{L}} \geq \varepsilon.$$

Пусть

$$N_{\mathcal{L}}(\varepsilon) = N_{\mathcal{L}}(\varepsilon, \mathcal{L}_*) = \sup_{B\subset\mathcal{L}_*} \#\{B\},$$

где супремум берётся по всем подмножествам $B$ компакта $\mathcal{L}_*$, состоящим из $\varepsilon$-различимых (в смысле метрики пространства $\mathcal{L}$) точек,

$$N_L(\varepsilon) = N_L(\varepsilon, \mathcal{L}_*) = \sup_{B\subset\mathcal{L}_*} \#\{B\},$$

где супремум берётся по всем подмножествам $B$ компакта $\mathcal{L}_*$, состоящим из $\varepsilon$-различимых (в смысле метрики пространства $\mathcal{L}$) точек.

Используя (14), можем написать следующее:

$$N_L(C_2^{-1}\varepsilon) \leq N_{\mathcal{L}}(\varepsilon) \leq N_L(C_1^{-1}\varepsilon), \tag{15}$$

где постоянные $C_1$, $C_2$ те же, что и в (14).

В работе Колмогорова А. Н., Тихомирова В. М. (1959) дается следующая асимптотика для величины $N_L(\varepsilon)$:

$$\ln N_L(\varepsilon) \asymp \left(\frac{1}{\varepsilon}\right)^{\frac{1}{\beta}}. \tag{16}$$

Отсюда и из (15) мы выводим справедливость следующей леммы.

**Лемма 2.** *Имеет место оценка*

$$\ln N_{\mathcal{L}}(\varepsilon) \geq C(\beta, \tau, C_{\mathcal{L}_*}) \left(\frac{1}{\varepsilon}\right)^{\frac{1}{\beta}}. \tag{17}$$

## 4. Доказательство теоремы 1

Обозначим

$$y[\varphi] = \int \varphi\, dy(t) = s[\varphi] + \sigma W[\varphi],$$

где

$$s[\varphi] = \int \varphi s\, dt, \quad W[\varphi] = \int \varphi\, dW(t).$$

Мы хотим оценить неизвестную функцию $s \in \mathcal{L}_*$,

$$s(t) = \sum_{u\in\Lambda} a(u)e^{iut}, \quad \sum_{u\in\Lambda} |a(u)|^2|u|^{2\beta} < C,$$

по наблюдениям

$$y[\varphi], \ \operatorname{supp}\varphi \subset [-T,T].$$

Обозначим $L^2([-T,T])$ – $L^2$–пространство на отрезке $[-T,T]$, построенное по нормированной мере Лебега:

$$(\psi_1,\psi_2)_{L^2_T} = \frac{1}{2T}\int_{-T}^{T} \psi_1\overline{\psi_2}\, dt.$$

Как известно (см. Н. Винер, Р. Пэли (1964)), если $T$ достаточно велико ($T \geq T_0(\tau)$), то

$$c_1\|s\|_{\mathcal{L}} \leq \|s\|_{L^2_T} \leq c_2\|s\|_{\mathcal{L}}, \ s \in \mathcal{L}(\Lambda), \tag{18}$$

где $c_1, c_2$ – константы, зависящие только от $\tau$. Примем обозначение $L^2_T$ для гильбертова пространства образованного из линейного множества $\mathcal{L}(\Lambda)$ с гильбертовой структурой, индуцированной пространством $L^2([-T,T])$.

Предположим, система $\{\psi_u,\, u \in \Lambda\}$ в $L^2_T$ выбрана так, что она является базисом Рисса в $L^2_T$. В этом случае наши наблюдения мы можем отождествить со случайным процессом $y(u)$, $u \in \Lambda$,

$$y(u) = \frac{1}{2T}\int_{T}^{T} \overline{\psi}\, dy(t) = a_s(u) + \sigma w(u). \tag{19}$$

Здесь

$$a_s(u) = (s,\psi_u)_{L^2_T}\,, \ w(u) = \frac{1}{2T}\int_{-T}^{T} \overline{\psi}\, dW(t). \tag{20}$$

Так что $y(u), u \in \Lambda$ – гауссовские величины со средним $a_s(u)$ и дисперсией $\sigma(u,u)$,

$$\sigma(u,u) = \frac{\sigma^2}{T} \times \frac{1}{2T} \int_{-T}^{T} |\psi_u|\, dt.$$

Очевидно, ковариационная матрица $\Sigma = (\sigma(u,v))$ случайного вектора $(y(u), u \in \Lambda)$ имеет вид

$$\sigma(u,v) = \frac{\sigma^2}{T} \times \frac{1}{2T} \int_{-T}^{T} \psi_u \overline{\psi_u}\, dt = \frac{\sigma^2}{T} \times r(u,v).$$

Обозначим $P_s$ распределение гауссовского вектора $(y_u, u \in \Lambda)$. Простые вычисления показывают, что

$$K(P_{s^1}, P_{s^2}) = \frac{T}{\sigma^2} \sum_{u,v \in \Lambda} \gamma(u,v)(a_{s^1}(u) - a_{s^2}(u))\overline{(a_{s^1}(u) - a_{s^2}(u))}, \quad (21)$$

где $\Gamma = (\gamma(u,v))$ – матрица, обратная к $R = (r(u,v))$.

Систему $\{\psi_u,\ u \in \Lambda\}$ в $L_T^2$ будем называть равномерно риссовским базисом, если равномерные нормы операторов с матрицами $R, \Gamma$ ограничены равномерно по $T$. Укажем, как выбрать такую систему.

Положим

$$\varphi_u(t) = e^{iut} \mathbb{1}_{[-T,T]}(t),\ u \in \Lambda.$$

Обозначим $\{\varphi_u^*(t),\ u \in \Lambda\}$ – сопряженную систему:

$$(\varphi_u, \varphi_v^*)_{L_T^2} = \delta_{u,v}, \quad (22)$$

где $\delta_{u,v}$ – символ Кронекера. Из (14), (13) следует, что система $\{\psi_u = \varphi_u^*(t),\ u \in \Lambda\}$ является равномерно риссовским базисом, если $T > T_0(\tau)$. Зафиксируем этот вывод в следующем утверждении.

**Лемма 3.** *Пусть выполнено условие* 12. *Тогда система*

$$\{\psi_u = \varphi_u^*(t),\ u \in \Lambda\}$$

*является равномерно риссовским базисом, если* $T > T_0(\tau)$ *и*

$$K(P_{s^1}, P_{s^2}) \le C(\tau) \frac{T}{\sigma^2} \sum_{u \in \Lambda} |a_{s^1}(u) - a_{s^2}(u))|^2. \quad (23)$$

Последнее неравенство следует из равномерной ограниченности оператора Г.

Теперь укажем каким образом будет выбрана система

$$\{s^1, \ldots, s^M\}$$

$\varepsilon-$различимых точек в $\mathcal{L}_*$. Мы возьмем

$$\varepsilon = \left(\frac{1}{2T}\right)^{\frac{\beta}{2\beta+1}}.$$

Указанную систему $\{s^1, \ldots, s^M\}$ мы выберем так, чтобы

$$\sigma^2 \sum_{u\in\Lambda} |a_{s^i}(u) - a_{s^j}(u))|^2 \leq \delta^2,$$

а величина $M$ оказалась достаточно большой:

$$1/2 \geq \frac{\frac{C(\tau)T}{\sigma^2}\delta^2 + \ln 2}{\ln(M-1)} \geq \frac{\sup_{i,j} K(P_{s^i}, P_{s^j}) + \ln 2}{\ln(M-1)}.$$

В этом случае, как это следует из леммы 1 мы получаем нужную оценку снизу для риска. Здесь мы использовали неравенство

$$K(P_{s^1}, P_{s^2}) \leq C(\tau)\frac{T}{\sigma^2}\delta^2,$$

верное для достаточно больших $T$, как это следует из леммы 3. Выбор такой системы $\{s^1, \ldots, s^M\}$ осуществляется так же, как это делается в [1] при доказательстве леммы 2.

## Литература

1. А. Н. Колмогоров, В. М. Тихомиров, *Эпсилон-энтропия и эпсилон-емкость множеств в функциональных пространствах.* — Успехи мат. наук **86** (2) (1959), 3–86.
2. Н. Винер, Р. Пэли, *Преобразование Фурье в комплексной области.* Наука, М. (1964).

Reshetov S. V. Lower bounds of risk at the problem of estimating of pseudo-periodic function, which is observed in white noise.

We consider the estimating problem, as we observe on large interval [-T,T] a pseudo-periodic function in white noise. We construct a lower minimax bound of the risk of all estimators.

Санкт-Петербургский
государственный университет

Поступило 8 декабря 2006 г.