Записки научных
семинаров ПОМИ
Том 339, 2006 г.

В. Н. Солев

ОЦЕНКИ МИНИМАЛЬНОГО РАССТОЯНИЯ

Введение

Пусть $X_1, \dots, X_n$ – независимые одинаково распределенные элементы измеримого пространства $(\mathcal{X}, \mathcal{B})$ с общим распределением $P \in \mathcal{P}$. Задача состоит в оценивании по наблюдениям $X_1, \dots, X_n$ неизвестной плотности $f = f(P)$,

$$f = f(P) = \frac{dP}{d\mu}. \tag{1}$$

Здесь $\mathcal{P}$ известное множество, мы предполагаем, что все распределения $P \in \mathcal{P}$ абсолютно непрерывны по отношению к σ-конечной мере μ и используем обозначение

$$\mathcal{F} = \{f : \ f = f(P), \text{ для некоторой } P \in \mathcal{P}\}. \tag{2}$$

$$\mathcal{F}_* = \{g : \ g = f_1 - f_2, \text{ для некоторых } f_1, f_2 \in \mathcal{F}\}. \tag{3}$$

Мы предполагаем, что $\mathcal{F} \subset L^2(d\mu)$ и обозначаем K наименьшее центрально-симметричное выпуклое множество, содержащее $\mathcal{F}_*$. Мы будем использовать следующую полунорму,порожденную K

$$\|\varphi\|_{\mathcal{P}} = \sup_{h \in \mathcal{F}_*} \int \varphi h \, d\mu = \sup_{h \in K} \int \varphi h \, d\mu. \tag{4}$$

Для двух вероятностных распределений μ_1 и μ_2 с плотностями f_1, f_2 по отношению к мере $\nu = \mu_1 + \mu_2$ мы используем обозначение $|\mu_1 - \mu_2|$ для меры

$$|\mu_1 - \mu_2| \{A\} = \int_A |f_1 - f_2| \, d\nu.$$

Для центрально-симметричного множества $\mathcal{C}$ функций φ и вероятностных мер P, Q мы используем обозначение

$$\kappa_{\mathcal{C}}(P, Q) = \sup_{\varphi \in \mathcal{C}} \int \varphi(x) d(P - Q). \tag{5}$$

Работа выполнена при поддержке грантов РФФИ 05-01-00920, НШ – 4222.2006.1, РФФИ–ННИО 04-01-04000.

Здесь мы предполагаем, что $\mathcal{C} \subset L^1(d|P-Q|)$. Здесь и далее интегралы без указания пределов берутся по всему пространству.

Если $P, Q \in \mathcal{P}$ мы будем писать

$$\kappa_{\mathcal{C}}(f(P), f(Q)) := \kappa_{\mathcal{C}}(P, Q).$$

Пусть P_n – эмпирическая распределение, построенное по выборке $X_1, \ldots, X_n$,

$$P_n\{A\} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}_A(X_j).$$

Здесь $\mathbb{I}_A(x)$ – индикатор множества A.

Оценку $\widehat{f}_n$ минимального расстояния для неизвестной плотности $f \in \mathcal{F}$ разумно определить соотношениями

$$\widehat{f}_n \in \mathcal{F}, \quad \kappa_{\mathcal{C}}(P_n, P_{\widehat{f}_n}) \le \kappa_{\mathcal{C}}(P_n, P_g), \text{ для } g \in \mathcal{F}. \tag{6}$$

Здесь и далее мы используем обозначение P_f для распределения из $\mathcal{P}$ с плотностью f. Данное выше определение продуктивно только в том случае, когда величина $\kappa_{\mathcal{C}}(P_n, \mathcal{P}) \to 0$ при $n \to \infty$. Очевидным препятствием для этого может быть то обстоятельство, что в большинстве случаев эмпирическая мера P_n и меры $P \in \mathcal{P}$ ортогональны. Простой выход из такого рода затруднения состоит в том, что взамен $\mathcal{C}$ мы выбираем подходящую конечную ε-сеть $\mathcal{C}_N$,

$$\begin{gathered} \mathcal{C}_N = \{\varphi_1, \ldots, \varphi_N\} \subset \mathcal{C}, \\ \sup_{\varphi \in \mathcal{C}} \inf_j \|\varphi - \varphi_j\|_{\mathcal{P}} \le \varepsilon. \end{gathered} \tag{7}$$

По некоторым техническим причинам такая замена удобна и в случае, когда $\kappa_{\mathcal{C}}(P_n, \mathcal{P}) \to 0$ при $n \to \infty$.

Итак, пусть $\mathcal{C}_N = \{\varphi_1, \ldots, \varphi_N\}$ конечное центрально-симметричное подмножество множества $\mathcal{C}$. Оценка $\widehat{f}_n$ минимального расстояния для неизвестной плотности $f \in \mathcal{F}$ определяется соотношениями

$$\widehat{f}_n \in \mathcal{F}, \quad \kappa_{\mathcal{C}_N}(P_n, P_{\widehat{f}_n}) \le \kappa_{\mathcal{C}_N}(P_n, P_g), \text{ для } g \in \mathcal{F}. \tag{8}$$

Исследуя свойства таких оценок мы выбираем ε-сеть $\mathcal{C}_N$ в $\mathcal{C}$, надлежащим образом связывая ε и n. При этом нашей целью является оценка величины минимаксного риска

$$\sup_{f \in \mathcal{F}} \mathbb{E}_f\, \kappa_{\mathcal{C}}(f, \widehat{f}_n).$$

Здесь и далее $\mathbb{E}_f$ – математическое ожидание, взятое по распределению P_f.

Два примера

Здесь и далее для $p \geq 1$ мы используем обозначение $L^p(d\mu)$ для L^p–пространства с нормой

$$\|\psi\|_p = \left\{ \int |\psi|^p \, d\mu \right\}^{1/p}.$$

Пример 1. Пусть

$$\mathcal{F}_* = \{g : \ g = f_1 - f_2 \text{ для некоторых } f_1, f_2 \in \mathcal{F}\}.$$

В случае, когда $\mathcal{F}_* \subset L^p(d\mu)$ мы возьмем

$$\mathcal{C}_p = \left\{ s : \ s = \operatorname{sign}(g) \frac{|g|^{p-1}}{\|g\|_p^{p-1}} \text{ для некоторой } g \in \mathcal{F}_* \right\}.$$

Здесь $p \geq 1$, $1/p + 1/q = 1$. Ясно, что, если $\mathcal{F}_* \subset L^p(d\mu)$, тогда множество $\mathcal{C}_p$ лежит на единичной сфере пространства $L^q(d\mu)$ и для любых $P, Q \in \mathcal{P}$,

$$\kappa_{\mathcal{C}_p}(P, Q) = \|f(P) - f(Q)\|_p \tag{9}$$

Пример 2. Пусть $(\mathcal{X}, \rho)$ – метрическое пространство. Введем в рассмотрение метрику Канторовича $\kappa(P, Q)$:

$$\kappa(P, Q) = \inf_{\nu} \iint\limits_{\mathcal{X} \times \mathcal{X}} \rho(x, y) \, \nu(dx, dy),$$

Здесь inf берется по всем вероятностным распределениям ν на $\mathcal{X} \times \mathcal{X}$ с маргиналами P и Q. Будем писать $\psi \in$ C-Lip, если

$$\|\psi\|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|\psi(x) - \psi(y)|}{\rho(x, y)} \leq C,$$

где sup берется по всем $x, y \in \mathcal{X}, x \neq y$.

Пусть

$$\mathcal{C} = \{\varphi : \ \|\varphi\|_{\mathrm{Lip}} \leq 1\}.$$

Тогда (см. Л. В. Канторович (1942), Л. В. Канторович, Г. Ш. Рубинштейн (1958))

$$\kappa_{\mathcal{C}}(P, Q) = \sup_{\varphi \in \mathcal{C}} \int \varphi(x) d(P - Q) = \kappa(P, Q).$$

Энтропия, двойственность метрической энтропии

Прежде всего опишем проблему, которая будет интересовать нас в этом пункте. Пусть $(\mathcal{X},\mathcal{B})$ – измеримое пространство, μ – сигма-конечная мера на $(\mathcal{X},\mathcal{B})$, $L^2(d\mu)$ – L^2-пространство с нормой $\|\cdot\|_2=\|\cdot\|_{L^2(d\mu)}$ и скалярным произведением

$$(\varphi,\psi)_{L^2(d\mu)}=\int\varphi\overline{\psi}\,d\mu.$$

Пусть $\mathcal{F}$ – семейство вероятностных плотностей относительно меры μ,

$$\mathcal{P}=\left\{P:\frac{dP}{d\mu}=f \text{ для некоторой } f\in\mathcal{F}\right\}.$$

Предположим, что $\mathcal{F}$ – подмножество пространства $L^2(d\mu)$, $\mathcal{C}$ центрально-симметричное семейство функций, $\mathcal{C}_N$ – конечное подмножество семейства $\mathcal{C}$, состоящее из N элементов. Возможно ли выбрать $\mathcal{C}_N$ таким образом, чтобы для любых $f_1,f_2\in\mathcal{F}$

$$\kappa_{\mathcal{C}}(f_1,f_2)=\sup_{\varphi\in\mathcal{C}}\int\varphi(f_1-f_2)\le\kappa_{\mathcal{C}_N}(f_1,f_2)+\varepsilon=\sup_{\varphi\in\mathcal{C}_N}\int\varphi(f_1-f_2)+\varepsilon.$$

Если такой выбор возможен, как связана величина $N=N(\varepsilon)$ с метрической энтропией множества $\mathcal{P}$? Ответ положителен, например, в случае, когда $\mathcal{F}$ – тотально ограниченное подмножество пространства $L^2(d\mu)$ и семейство $\mathcal{C}$ помещается в шар конечного радиуса.

Чтобы сформулировать точный результат, мы напомним некоторые обозначения. Мы обозначили K наименьшее выпуклое центрально-симметричное тело, содержащее $\mathcal{F}_*=\mathcal{F}-\mathcal{F}$,

$$\mathcal{F}_*=\{g:g=f_1-f_2\quad\text{для некоторых}\quad f_1,f_2\in\mathcal{F},\}$$

и ввели полунорму

$$\|\varphi\|_{\mathcal{P}}=\sup_{h\in K}\int\varphi h\,d\mu. \tag{10}$$

Пусть K^0 полярное множество для K:

$$K^0=\left\{h:\sup_{x\in K}(h,x)_{L^2(d\mu)}=\|h\|_{\mathcal{P}}\le 1\right\}.$$

Ясно, что, если $S \subset \left\{x + \varepsilon K^0\right\}$, то для любых $s_1, s_2 \in S$ и любого $f \in \mathcal{F}_*$

$$|\int (s_1 - s_2) f \, d\mu| \leq 2\varepsilon.$$

Таким образом, если возможно покрыть семейство $\mathcal{C}$ конечным числом сдвигов множества εK^0:

$$\mathcal{C} \subset \bigcup_{j=1}^{N} \left\{x_j + \varepsilon K^0\right\},$$

то для $f_1, f_2 \in \mathcal{F}$ получаем

$$\kappa_{\mathcal{C}}(f_1, f_2) \leq \kappa_{\mathcal{C}_N}(f_1, f_2) + 2\varepsilon, \tag{11}$$

где

$$\mathcal{C}_N = \left\{y_1, \ldots, y_N\right\}, \;\; y_j \in \left\{x_j + \varepsilon K^0\right\} \cap \mathcal{C}.$$

В метрическом пространстве $(\mathcal{Y}, \rho)$ мы обозначаем $V_\varepsilon(h)$ шар радиуса ε с центром в точке h. Для подмножества $\mathcal{D} \subset \mathcal{Y}$ обозначим $N = N(\varepsilon) = N(\varepsilon, \mathcal{D}, \rho)$ наименьшее N такое, что

$$\mathcal{D} \subset \bigcup_{j=1}^{N} V_\varepsilon(h_j), \;\; \text{где } h_j \in \mathcal{D}, \; j = 1, \ldots, N.$$

Величина $H(\varepsilon) = H(\varepsilon, \mathcal{D}, \rho) = \log N(\varepsilon, \mathcal{D}, \rho)$ называется метрической энтропией $\mathcal{D}$.

Если $\mathcal{Y}$ – Банахово пространство с нормой $\|\cdot\|$ и

$$\rho(x, y) = \|x - y\|,$$

мы будем писать $N(\varepsilon, \mathcal{D}, \|\cdot\|)$ и $H(\varepsilon, \mathcal{D}, \|\cdot\|)$ вместо $N(\varepsilon, \mathcal{D}, \rho)$ и $H(\varepsilon, \mathcal{D}, \rho)$.

Мы хотели бы сравнить энтропию $H(\varepsilon, \mathcal{C}, \|\cdot\|_{\mathcal{P}})$ и энтропию $H(\varepsilon, \mathcal{P}, \|\cdot\|_{L^2(d\mu)})$.

Предложение 1. Пусть K наименьшее центрально-симметричное выпуклое подмножество пространства $L^2(d\mu)$, содержащее семейство плотностей $\mathcal{F}$, $\mathcal{C}$ – центрально-симметричное подмножество, помещающееся в единичный шар пространства $L^2(d\mu)$. Тогда существуют такие универсальные константы $c > 0$ и $C > 0$, что для всех n, $\varepsilon > 0$

$$H(c\varepsilon, \mathcal{C}, \|\cdot\|_{\mathcal{P}}) \leq CH\left(\varepsilon, K, \|\cdot\|_{L^2(d\mu)}\right). \tag{12}$$

Прежде, чем перейти к доказательству предложения 1, приведем формулировку одного результата S. Artstein, V. Milman, S. J. Szarek, следствием которого предложение 1 является.

Для двух выпуклых тел K и T из $\mathbb{R}^n$, величина $N(K,T)$ определяется как число сдвигов T необходимое для того, чтобы покрыть K:

$$N(K,T)=\min\left\{N : \exists x_1,\dots,x_N\in\mathbb{R}^n,\ K\subset\bigcup_{j=1}^{n}\{T+x_j\}\right\}.$$

Обозначим K^0 полярное множество для K,

$$K^0=\left\{x : \sup_{y\in K}(x,y)\le 1\right\}.$$

Здесь (x,y) – скалярное произведение в $\mathbb{R}^n$. Пусть V – единичный шар в $\mathbb{R}^n$ с центром в начале координат.

Теорема (S. Artstein, V. Milman, S. J. Szarek (2004)). *Существуют две универсальные константы α и β такие, что для любого n, любого $\varepsilon>0$ и любого центрально-симметричного выпуклого подмножества $K\in\mathbb{R}^n$ имеют место неравенства*

$$N\left(V,\frac{\varepsilon}{\alpha}K^0\right)^{1/\beta}\le N\left(K,\varepsilon V\right)\le N\left(V,\alpha\varepsilon K^0\right)^{\beta}. \tag{14}$$

Доказательство предложения 1. Пусть

$$K=\{\psi : \psi=p_1f_1+p_2f_2 \text{ для некоторых } f_1,f_2\in\mathcal{F} \text{ и } |p_1|+|p_2|=1\},$$

$$K\subset\bigcup_{j=1}^{N}\{x_j+\varepsilon V\}. \tag{15}$$

Здесь V единичный шар пространства $L^2(d\mu)$ с центром в точке $\mathbf{0}$, элементы $x_1,\dots,x_N\in K$, предполагается, что покрытие множества K выбрано оптимальным образом: $N=N(\varepsilon,K,\|\cdot\|_{L^2(d\mu)})$.

Обозначим L_N – подпространство пространства $L^2(d\mu)$, порожденное элементами $x_1,\dots,x_N$, K_N – проекцию K на L_N, V_N – единичный шар подпространства L_N с центром в точке $\mathbf{0}$. Очевидно,

$$K_N\subset\bigcup_{j=1}^{N}\{x_j+\varepsilon V_N\},\quad N(\varepsilon,K_N,\|\cdot\|_{L^2(d\mu)})\le N. \tag{16}$$

Обозначим K_N^0 – полярное множество для K_N в L_N:

$$K_N^0 = \left\{ h : \sup_{x \in K_N} |(h,x)_{L^2(d\mu)}| \le 1, h \in L_N \right\}.$$

Из упомянутого выше результата S. Artstein, V. Milman, S. J. Szarek (2004) следует, что существуют две такие универсальные константы $d \ge 1$ и $C \ge 1$, такие точки $y_1, \ldots, y_{N_*}$ из L_N, что

$$V_N \subset \bigcup_{j=1}^{N_*} \left\{ y_j + d\,\varepsilon K_N^0 \right\}, \tag{17}$$

$$N_* \le N^C. \tag{18}$$

Обозначим L^N – ортогональное дополнение к L_N, V^N – единичный шар подпространства L^N с центром в точке $\mathbf{0}$. Заметим, что

$$|(h,x)_{L^2(d\mu)}| \le \varepsilon, \text{ если } x \in K,\ h \in L^N, ||h||_{L^2(d\mu)} \le 1. \tag{19}$$

Поэтому $V^N \subset \varepsilon K^0$ и $d\varepsilon K_N^0 + V_N \subset 2d\varepsilon K^0$. Очевидно, $V \subset V_N + V^N$ и, следовательно,

$$V \subset \bigcup_{j=1}^{N_*} \left\{ y_j + d\,\varepsilon K_N^0 + V^N \right\}. \tag{20}$$

Отсюда,

$$V \subset \bigcup_{j=1}^{N_*} \left\{ y_j + 2d\,\varepsilon K^0 \right\}. \tag{21}$$

Последнее вместе с (18) дает (12) при $c = 2d$.

Оценка минимального расстояния

Пусть $\mathcal{C}$ центрально-симметричное подмножество пространства $L^2(d\mu)$, которое помещается в шар $V_R(\mathbf{0})$ радиуса R с центром в точке $\mathbf{0}$. Далее, пусть выбрано центрально-симметричное подмножество $\mathcal{C}_N \subset \mathcal{C}$, состоящее из N элементов, такое, что

$$N \le 2N(\mathcal{C}, \varepsilon, ||\cdot||_{\mathcal{P}}), \tag{22}$$

$$\sup_{x \in \mathcal{C}} \inf_{y \in \mathcal{C}_N} ||x - y||_{\mathcal{P}} \le 2\varepsilon. \tag{23}$$

Пусть $X_1, \dots, X_n$ независимые одинаково распределенные случайные элементы со значениями в $(\mathcal{X}, \mathcal{B})$ с общим распределением $P \in \mathcal{P}$ и плотностью $f \in \mathcal{F}$. Мы хотим построить по выборке $X_1, \dots, X_n$ оценку $\widehat{f}_n \in \mathcal{F}$ для неизвестной плотности f.

Оценка минимального расстояния $\widehat{f}_n = \widehat{f}_n^N$ это плотность из $\mathcal{F}$, определяемая соотношением

$$\kappa_{\mathcal{C}_N}\left(P_n, P_{\widehat{f}_n}\right) \le \kappa_{\mathcal{C}_N}\left(P_n, P_g\right) \text{ для } \quad g \in \mathcal{F}. \tag{24}$$

Здесь P_g – распределение с плотностью g, P_n – эмпирическое распределение.

Лемма 1. *Если $f \in \mathcal{F}$, то*

$$\begin{aligned} &\kappa_{\mathcal{C}}\left(\widehat{f}_n, f\right) \\ &\le 2\varepsilon + \kappa_{\mathcal{C}_N}\left(\widehat{f}_n, f\right) = 2\varepsilon + \kappa_{\mathcal{C}_N}\left(P_{\widehat{f}_n}, P_f\right) \\ &\le 2\varepsilon + \kappa_{\mathcal{C}_N}\left(P_{\widehat{f}_n}, P_n\right) + \kappa_{\mathcal{C}_N}\left(P_f, P_n\right) \\ &\le 2\varepsilon + 2\kappa_{\mathcal{C}_N}\left(P_f, P_n\right). \end{aligned} \tag{25}$$

Таким образом, для того, чтобы получить информацию о величине $\kappa_{\mathcal{C}}\left(\widehat{f}_n, f\right)$ для истинного значения плотности f достаточно изучить асимптотическое поведение супремума эмпирического процесса

$$\xi(\varphi) = (P_n - P)\varphi = \int \varphi\, d(P_n - P), \quad \varphi \in \mathcal{C}_N.$$

Эмпирическиий процесс

Мы должны наложить некоторые условия на центрально-симметричное подмножество $\mathcal{C}$ и параметрическое множество $\mathcal{P}$. Далее мы предполагаем, что

$$\begin{gathered} \mathcal{C} \subset V_R(\mathbf{0}), \quad \text{где} \quad V_R(\mathbf{0}) \quad \text{шар радиуса} \quad R \\ \text{с центром в точке} \quad \mathbf{0} \quad \text{в пространстве} \quad L^2(d\mu). \end{gathered} \tag{26}$$

Мы предполагаем также что существует такая полу-метрика ψ, что для всех $x, y \in \mathcal{X}$, $\varphi \in \mathcal{C}$, $P \in \mathcal{P}$

$$|\varphi(x) - \varphi(y)| \le \psi(x, y), \tag{27}$$

$$\iint \psi^2(x,y)P(dx)P(dy) \le L = L(\mathcal{P},\mathcal{C}). \tag{28}$$

Обозначим

$$\vartheta_n = \sup_{\varphi\in\mathcal{C}_N} \xi(\varphi) = \kappa_{\mathcal{C}_N}(P_n, P).$$

Оценка величины $\mathbb{E}_P\vartheta_n$ производится обычным методом (см.L. Devroye, G. Lugoshi (2001)).

Рассмотрим эмпирический процесс

$$\xi(\varphi) = \frac{1}{n}\sum_{j=1}^{n}\left(\varphi(X_j) - \mathbf{E}_P\varphi(X_j)\right), \varphi \in \mathcal{C}_N.$$

Таким образом,

$$\mathbf{E}\vartheta_n = \mathbf{E}\sup_{\varphi\in\mathcal{C}_N}\frac{1}{n}\,|\sum_{j=1}^{n}\left(\varphi(X_j) - \mathbf{E}_P\varphi(X_j)\right)| \tag{29}$$

Пусть $X_1^\bullet,\dots,X_n^\bullet$ – независимые копии величин $X_1,\dots,X_n$. Хорошо известно, что

$$\mathbb{E}\vartheta_n = \mathbb{E}\sup_{\varphi\in\mathcal{C}_N}\left|\mathbb{E}\left\{\frac{1}{n}\sum_{j=1}^{n}\left(\varphi(X_j)-\varphi(X_j^\bullet)\right)|X_1,\dots,X_n\right\}\right|$$

$$\le \mathbb{E}\big(\sup_{\varphi\in\mathcal{C}_N}|\frac{1}{n}\sum_{j=1}^{n}\left(\varphi(X_j)-\varphi(X_j^\bullet)\right)|\big)$$

Таким образом,

$$\mathbb{E}\vartheta_n \le \mathbb{E}\big(\sup_{\varphi\in\mathcal{C}_N}|\frac{1}{n}\sum_{j=1}^{n}\left(\varphi(X_j)-\varphi(X_j^\bullet)\right)|\big). \tag{30}$$

Далее рассмотрим n независимых одинаково распределенных радемахеровских величин $\varepsilon_1,\dots,\varepsilon_n$ $P\{\varepsilon_1=1\} = P\{\varepsilon_1=-1\} = 1/2$, независимых от $X_1,\dots,X_n,X_1^\bullet,\dots,X_n^\bullet$. Ясно, что

$$\mathbb{E}\big(\sup_{\varphi\in\mathcal{C}_N}|\frac{1}{n}\sum_{j=1}^{n}\left(\varphi(X_j)-\varphi(X_j^\bullet)\right)|\big)$$

$$= \mathbb{E}\big(\sup_{\varphi\in\mathcal{C}_N}|\frac{1}{n}\sum_{j=1}^{n}\varepsilon_j\left(\varphi(X_j)-\varphi(X_j^\bullet)\right)|\big), \tag{31}$$

поскольку случайные величины $\left(\varphi(X_j - \varphi(X_j^\bullet)\right)$ и $\varepsilon_j\left(\varphi(X_j) - \varphi(X_j^\bullet)\right)$ одинаково распределены.

На следующем шаге мы используем хорошо известную лемму (см.P. Massart (2000)).

Лемма 2. *Пусть A – конечное подмножество пространства $\mathbb{R}^n$, состоящее из N элементов, $\varepsilon_1, \dots, \varepsilon_n$ радемахеровские величины, $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, $R = \sup_{a\in A} \|a\| = \sup_{a\in A} \sqrt{\sum_{j=1}^n a_j^2}$. Тогда*

$$\mathbf{E} \sup_{a\in A} \sum_{j=1}^n a_j \varepsilon_j \le R\sqrt{2\ln N}. \tag{32}$$

В нашем случае $na_j = \varphi(X_j) - \varphi(X_j^\bullet)$ – случайные величины, независимые от $\varepsilon_1, \dots, \varepsilon_n$. Поэтому при условии (27)

$$R^2 = \sup_{a\in A} \sum_{j=1}^n a_j^2 \le \frac{1}{n^2} \sum_{j=1}^n \psi^2(X_j, X_j^\bullet).$$

Отсюда, используя (32), мы выводим следующую лемму.

Лемма 3. *При условиях (27), (28)*

$$\mathbf{E}_P \kappa_{\mathcal{C}_N}(P_n, P) \le \sqrt{L(\mathcal{P}, \mathcal{C})} \sqrt{\frac{2\ln|\mathcal{C}_N|}{n}}. \tag{33}$$

Здесь для множества A мы обозначили через $|A|$ число элементов A.

Основной результат

Пусть $X_1, \dots, X_n$ – независимые одинаково распределенные элементы измеримого пространства $(\mathcal{X}, \mathcal{B})$ с общим распределением $P \in \mathcal{P}$ и плотностью $f \in \mathcal{F}$ по отношению к σ-конечной мере μ.

Пусть, далее, $\mathcal{C}$ центрально-симметричное подмножество пространства $L^2(d\mu)$. Пусть выбрано центрально-симметричное подмножество $\mathcal{C}_N \subset \mathcal{C}$, состоящее из N элементов, а оценка минимального расстояния $\widehat{f}_n = \widehat{f}_n^N$ это плотность из $\mathcal{F}$, определяемая соотношением

$$\kappa_{\mathcal{C}_N}\left(P_n, P_{\widehat{f}_n}\right) \le \kappa_{\mathcal{C}_N}\left(P_n, P_g\right) \text{ для} g \in \mathcal{F}. \tag{34}$$

Теорема. *Предположим, что параметрическое множество $\mathcal{F}$ помещается в выпуклое центрально-симметричное подмножество $K \subset L^2(d\mu)$ такое, что*

$$H\left(\varepsilon, K, \|\cdot\|_{L^2(d\mu)}\right) \le C\left(\frac{1}{\varepsilon}\right)^{\frac{1}{\gamma}}, \tag{35}$$

центрально-симметричное подмножество $\mathcal{C}$ помещается в шар $V_R(\mathbf{0})$ радиуса R с центром в точке $\mathbf{0}$. Тогда при условиях (27), (28) центрально-симметричное подмножество $\mathcal{C}_N \subset \mathcal{C}$, состоящее из N элементов может быть выбрано так, чтобы риск оценки минимального расстояния $\widehat{f}_n$ удовлетворял неравенству

$$\sup_{P\in\mathcal{P}} \mathbb{E}_P\, \kappa_{\mathcal{C}}(\widehat{f}_n, f) \le \mathcal{K}(\mathcal{P},\mathcal{C})\left(\frac{1}{n}\right)^{\frac{\gamma}{2\gamma+1}} \tag{36}$$

Доказательство. Мы можем предполагать, что K – выпуклая оболочка множества $\mathcal{F}_* = \mathcal{F} - \mathcal{F}$. По предложению 1 существуют такие абсолютные константы c, C, что для любого $\varepsilon > 0$

$$H\left(cR\varepsilon, \mathcal{C}, \|\cdot\|_{\mathcal{P}}\right) \le CH\left(\varepsilon, K, \|\cdot\|_{L^2(d\mu)}\right). \tag{37}$$

Отсюда и из (35) получаем

$$H\left(\varepsilon, \mathcal{C}, \|\cdot\|_{\mathcal{P}}\right) \le C_R\left(\frac{1}{\varepsilon}\right)^{\frac{1}{\gamma}}. \tag{38}$$

Поэтому центрально-симметричное подмножество $\mathcal{C}_N \subset \mathcal{C}$, состоящее из N элементов может быть выбрано так, чтобы

$$N \le 2N(\mathcal{C}, \varepsilon, \|\cdot\|_{\mathcal{P}}), \tag{39}$$

$$\sup_{x\in\mathcal{C}} \inf_{y\in\mathcal{C}_N} \|x-y\|_{\mathcal{P}} \le 2\varepsilon, \tag{40}$$

Как это следует из леммы 1

$$\kappa_{\mathcal{C}}\left(\widehat{f}_n, f\right) \le 2\varepsilon + 2\kappa_{\mathcal{C}_N}\left(P_f, P_n\right). \tag{41}$$

Выберем

$$\varepsilon = \varepsilon(n) = \left(\frac{1}{n}\right)^{\frac{\gamma}{2\gamma+1}}. \tag{42}$$

При таком выборе нам остается доказать, что

$$\mathbb{E}_f\, \kappa_{\mathcal{C}_N}\left(P_f, P_n\right) \le C(\mathcal{P},\mathcal{C})\left(\frac{1}{n}\right)^{\frac{\gamma}{2\gamma+1}}.$$

Последнее следует из леммы 3 и оценки (35).

Литература

1. Л. В. Канторович, *Перемещение масс.* — ДАН СССР **37**, (1942), 227–229.
2. Л. В. Канторович, Г. Ш. Рубинштейн, *О пространстве тотально аддитивных функций.* — Вестник ЛГУ **13**, (7) (1958), 52–59.
3. А. Н. Колмогоров, В. М. Тихомиров, *Эпсилон-энтропия и эпсилон-емкость множеств в функциональных пространствах.* — Успехи Мат. Наук **86**, (2) (1959), 3–86.
4. L. Devroye and G. Lugoshi, *Combinatorial methods in density estimation.* Springer-Verlag, New York (2001).
5. P. Massart, *Some applications of concentration inequality.* — Annales de la faculté des science de Toulouse, 6 serie, **9**, (2) (2000) 245–303.
6. L. Birgé, *On estimating of a density using Hellinger distance and Some others strange facts.* — Probab. Th. Rel. Fields **71** (1993), 271–293.
7. P. Massart, *Concentration Inequality and Model Selection.* — Ecole d'Eté de Probabilités de Saint-Flour XXXIII–2003, Springer, New York (2003), 1–335.
8. S. Artstein, V. Milman, and S. J. Szarek, *Duality of metric entropy.* (2004).
9. S. McDiarmid, *On the method of bounded differences.* — In: Surveys in Combinatorics. Cambridge University Press (1989) 148–188.

Solev V. N. Minimum distance estimators.

In the paper the estimation problem for unknown density on independent observations is considered. We use the minimum distance estimation method. It is shown, that the accuracy of estimating is connected to the rate of increase of the entropy of parametrical set.

С.-Петербургское отделение
Математического института
им. В. А Стеклова РАН

Поступило 28ноября 2006 г.