

А. А. ТЕМПЕЛЬМАН

К ОБЩЕЙ ТЕОРИИ ЛИНЕЙНЫХ ОЦЕНОК СРЕДНЕГО ЗНАЧЕНИЯ

(Представлено академиком А. Н. Колмогоровым 29 III 1974)

Заметка посвящена перенесению результатов работ ⁽⁴⁾ и ⁽¹⁰⁾ на нелинейные регрессионные модели. В п. 1 на такие модели распространяется понятие M -мажоранты ядра, введенное в этих работах; это позволяет почти дословно перенести на общий случай конструкцию несмещенных линейных оценок, а также формулировки и доказательства основных результатов из ⁽⁴⁾, ⁽¹⁰⁾ (см. ниже теоремы 1–4 в п. 2).

1. Пусть $(\Omega, \mathcal{F})$ — измеримое пространство; всякую измеримую вещественную функцию $\xi = \xi(\omega)$, $\omega \in \Omega$, будем называть случайной величиной (с.в.). В дальнейшем T — произвольное множество и $\xi(\omega, t)$, $t \in T$, — фиксированная случайная функция (с.ф.) на T . Распределение с.в. γ определяется заданием вероятностной меры (в.м.) на $\mathcal{F}$: если P — такая мера, то $E_P \gamma = \int_{\Omega} \gamma(\omega) P(d\omega)$, $D_P \gamma = E_P(\gamma - E_P \gamma)^2$; $L^2(\Omega, \mathcal{F}, P)$ — гильбертово пространство классов P -эквивалентных с.в. γ с $E_P \gamma^2 < \infty$, причем скалярное произведение $\langle \gamma_1, \gamma_2 \rangle_{L^2} = E_P \gamma_1 \gamma_2$; $\mathcal{P}$ — класс всех в.м. P , для которых $E_P[\xi(t)]^2 < \infty$, $t \in T$. Если $P \in \mathcal{P}$, то $H_P(\xi)$ — подпространство в $L^2(\Omega, \mathcal{F}, P)$, порожденное семейством с.в. $\{\xi(t), t \in T\}$.

Пусть $K(s, t)$ — ядро (неотрицательно определенная функция) на $T \times T$, $m(\cdot)$ — произвольная функция на T ; тогда $\mathcal{P}_K$ — класс всех в.м. P , для которых $E_P \xi(s) \xi(t) = K(s, t)$; $\mathcal{P}_m$ — класс всех в.м. P , для которых $E_P \xi(t) = m(t)$ и $E_P(\xi(s) - m(s))(\xi(t) - m(t)) = K(s, t)$; $H(K)$ — гильбертово пространство функций на T с воспроизводящим ядром K (см. ⁽⁶⁾).

Сеть множеств $\{a_i^\alpha, \dots, a_{n_\alpha}^\alpha; t_i^\alpha, \dots, t_{n_\alpha}^\alpha\}$, $\alpha \in A$, где A — направленное множество, а a_i^α — вещественные числа, $t_i^\alpha \in T$ и n_α — натуральные числа, назовем линейной операцией, определяющей элемент $\gamma \in H_P(\xi)$

(соответственно элемент $j \in H(K)$), если $\gamma = (P) \text{Li.m.} \sum_{\alpha \in A} a_i^\alpha \xi(t_i^\alpha)$ (соответ-

ственно $f(\cdot) = \lim_{\alpha \in A} \sum_{i=1}^{n_\alpha} a_i^\alpha R(\cdot, t_i^\alpha)$ в $H(K)$); класс всех таких операций обозна-

чим $LO(\gamma)$ (соответственно $LO(f)$); $LO(P) = \bigcup_{\gamma \in H_P(\xi)} LO(\gamma)$; $LO(K) =$

$= \bigcup_{f \in H(K)} LO(f)$. Пусть $P, Q \in \mathcal{P}$: элемент $\gamma' \in H_P(\xi)$ называется линейным

образом элемента $\gamma \in H_Q(\xi)$, если $LO(\gamma) \subset LO(\gamma')$; элементы γ и γ' вполне подобны, если $LO(\gamma) = LO(\gamma')$; элементы γ и γ' подобны относительно класса линейных операций $W \subset LO(P) \cap LO(Q)$, если $LO(\gamma) \cap W = LO(\gamma') \cap W \neq \emptyset$. Будем говорить, что ядро K мажорируется ядром B ($K \triangleleft B$), если $B - K$ — ядро; K ограничено относительно B ($K \triangleleft B$), если $K \triangleleft \lambda B$ при некотором $\lambda > 0$. Условие $K \triangleleft B$ необходимо и достаточно, для того чтобы любой элемент $\gamma \in H_Q(\xi)$, $Q \in \mathcal{P}_B$, обладал линейным образом в $H_P(\xi)$, $P \in \mathcal{P}_K$.

Напомним, что $K \triangleleft \lambda B$ равносильно тому, что $K(\cdot, t) \in H(B)$ при $t \in T$ и оператор $L_K^B: (L_K^B f)(t) = \langle K(\cdot, t), f(\cdot) \rangle_{H(B)}$ в $H(B)$ имеет норму $\|L_K^B\| \leq$

$\leq \lambda$; если K и B — обобщенные стационарные ядра на вещественном m -мерном евклидовом пространстве E^m со спектральными мерами F_K и F_B , то $K \triangleleft \lambda B$ равносильно тому, что $F_K \leq F_B$ и $|dF_K/dF_B| \leq \lambda$ F_B -почти всюду на E^m .

Согласно известной теореме Лозва, существует единственный изометрический изоморфизм i между пространствами $H(K)$ и $H_P(\xi)$, $P \in \mathcal{P}_K$,

такой, что $K(\cdot, t) \leftrightarrow \xi(t)$; элемент $\gamma = i(f) \in H_P(\xi)$ мы будем также обозначать: $\langle \xi, f \rangle_{H(K)}$; символами γ и $\langle \xi, f \rangle_{H(K)}$ будем обозначать и линейные образы этого элемента. Таким образом, если $j \in H(B)$, символ $\langle \xi, f \rangle_{H(B)}$ определен и в пространствах $H_P(\xi)$, $P \in \mathcal{P}_K$ с $K \triangleleft B$; при этом оператор $f \mapsto \langle \xi, f \rangle_{H(B)}$ является непрерывным линейным оператором из $H(B)$ в $H_P(\xi)$ и $B(\cdot, t) \mapsto \xi(t)$. Если $m(\cdot) \in H(B)$ и $R \triangleleft B$, то и $K(s, t) = R(s, t) + m(s)m(t) \triangleleft B(s, t)$ и, следовательно, при любом $j \in H(B)$ в любом пространстве $H_P(\xi)$, $P \in \mathcal{P}_{R, m}$, определен элемент $\langle \xi, f \rangle_{H(B)}$; если $\xi(\cdot) \in H(B)$ с P -вероятностью 1, то с.в. $\langle \xi, f \rangle_{H(B)}$ с P -вероятностью 1 совпадает со скалярным произведением $\xi(\cdot)$ и $f(\cdot)$ в $H(B)$.

Пусть R — ядро и M — произвольное семейство функций на T , содержащееся в некотором гильбертовом пространстве $H(\Phi)$ с воспроизводящим ядром Φ . Тогда для любого ядра R существует по крайней мере одно ядро $\tilde{R}$, такое, что: 1) $R \triangleleft \tilde{R}$, 2) $M \subset H(\tilde{R})$, 3) $H(\tilde{R} - R) \equiv \bar{M}(\tilde{R})$, где $\bar{M}(\tilde{R})$ — подпространство в $H(\tilde{R})$, порожденное семейством M . Всякое такое ядро мы будем называть M -мажорантой ядра R . Если M — конечномерное линейное пространство, то класс всех M -мажорант ядра R

совпадает с классом всех функций вида $\tilde{R}(s, t) = R(s, t) + \sum_{i=1}^n b_i \varphi_i(s) \varphi_i(t)$

($\{b_{ij}\}$ — неотрицательно определенная матрица и $\varphi_i \in M$) таких, что $M \subset H(\tilde{R})$; это условие выполнено, если $\det \{b_{ij}\} > 0$ и $\{\varphi_i\}$ — базис в M (подробнее см. ⁽¹⁰⁾).

2. Пусть известно, что $P \in \mathcal{P}_{R, M} = \bigcup_{m \in M} \mathcal{P}_{R, m}$, и по наблюдаемой реализа-

ции с.ф. $\xi(t)$, $t \in T$, над вероятностным пространством $(\Omega, \mathcal{F}, P)$ требуется оценить значение $\psi(m)$ определенного на M функционала ψ при $m(t) \equiv E_P \xi(t)$. Пусть $\mathcal{P}_0 \subset \mathcal{P}$; $\mathcal{P}_0$ -линейной статистикой класса $W \subset LO(P)$, построенной по реализации с.ф. $\xi(t)$ на T , назовем произвольное семейство подобных относительно класса W элементов $\{\psi_P, P \in \mathcal{P}_0\}$. Подобие элементов ψ_P позволяет рассматривать статистику $\{\psi_P\}$ как линейный функционал на реализациях $\xi(t)$, $t \in T$, который строится «операционно независимо» от неизвестной меры P^* . В работе ⁽¹¹⁾ для $\mathcal{P}_0 = \mathcal{P}_{R, M}$ введено более сильное условие операционной независимости гильбертовых пространств $(LO(P) = LO(Q))$ при любых $P, Q \in \mathcal{P}_{R, M}$ и показано, что это условие выполнено, если $M \subset H(R)$, а также в некоторых случаях при $M \not\subset H(R)$. Отметим, что это условие выполнено, если $P, Q \in \mathcal{P}_{R, M}$, каково бы ни было M .

$\mathcal{P}_{R, M}$ -линейная статистика $\hat{\psi}$ называется $\mathcal{P}_{R, M}$ -несмещенной линейной оценкой значения функционала $\psi(m)$ при $m(t) \equiv E_P \xi(t)$ ($\mathcal{P}_{R, M}$ -н.л.о.), если $E_P \hat{\psi} = \psi(m)$, коль скоро $P \in \mathcal{P}_{R, M}$, $m \in M$; $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}^0$ называется наилучшей, если $D_P \hat{\psi}^0 \leq D_P \hat{\psi}$ при всех $P \in \mathcal{P}_{R, M}$ для любой $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}$. Функционал ψ $\mathcal{P}_{R, M}$ -оцениваем с помощью операций класса W , если для него существует хотя бы одна $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}$ класса W . В дальнейшем в качестве W часто рассматриваются классы $LO(\tilde{R}) \equiv \equiv LO(R)$, где $\tilde{R}$ — M -мажоранты ядра R . При этом, если $M \subset H(R)$ и $\tilde{R} = R$, элементы ψ_P вполне подобны один другому; в противном случае они являются линейными образами вполне подобных элементов $\hat{\psi}_Q$, образуя-

* В дальнейшем мы будем писать ψ вместо ψ_P .

щих $\mathcal{P}_{R, M}$ -н.л.о.; если M — конечномерное линейное пространство, класс $LO(\tilde{R})$ не зависит от выбора $\tilde{R}$ *.

Обозначим $G_\psi(\tilde{R}) = \{g: g \in H(\tilde{R}), \psi(m) = \langle m, g \rangle_{H(\tilde{R})} \text{ на } M\}$; $\Pi_{\tilde{M}(\tilde{R})}^{H(\tilde{R})}$ — оператор проектирования в $H(\tilde{R})$ на подпространство $\tilde{M}(\tilde{R})$, порожденное множеством M . Легко убедиться, что если $G_\psi(\tilde{R}) \neq \emptyset$, то элемент $\pi_\psi(\tilde{R}) = \Pi_{\tilde{M}(\tilde{R})}^{H(\tilde{R})} g$ не зависит от выбора функции g из $G_\psi(\tilde{R})$ и является (единственным) элементом множества $G_\psi \cap \tilde{M}(\tilde{R})$; таким образом, условие $G_\psi(\tilde{R}) \neq \emptyset$ равносильно тому, что функционал ψ может быть продолжен до непрерывного линейного функционала на $\tilde{M}(\tilde{R})$.

Следующая теорема обобщает результаты Парзена ⁽⁷⁾ и автора ^(4, 10).

Теорема 1. *Функционал $\psi(m)$, $m \in M$, $\mathcal{P}_{R, M}$ -оцениваем с помощью операций из $LO(\tilde{R})$ тогда и только тогда, когда $G_\psi(\tilde{R}) \neq \emptyset$; при этом статистика $\hat{\psi}_{R, M} = \langle \xi, \pi_\psi(\tilde{R}) \rangle_{H(\tilde{R})}$ является единственной наилучшей $\mathcal{P}_{R, M}$ -н.л.о.*

К числу функционалов, $\mathcal{P}_{R, M}$ -оцениваемых с помощью операций из $LO(\tilde{R})$ при любой M -мажоранте $\tilde{R}$, принадлежат все функционалы $\psi_i(m) = m(t)$, $t \in T$, так как $\tilde{R}(\cdot, t) \in G_{\psi_t}(\tilde{R})$.

Будем говорить, что $\mathcal{P}_{R, M}$ -н.л.о. ψ определяется ядром B , если при любой в.м. $P \in \mathcal{P}_{R, M}$ с P -вероятностью 1 $\hat{\psi} = \hat{\psi}_{B, M}$ (т. е. $\hat{\psi}$ — линейный образ в $H_P(\xi)$ элемента $\hat{\psi}_{B, M} \in H_Q(\xi)$, $Q \in \mathcal{P}_{B, M}$) ** (см. также ^(4, 10)). Понятие оценки, определяемой ядром, близко к введенному ранее Ю. А. Розановым ^(2, 3) понятию псевдонаилучшей оценки.

Теорема 2. *Если $R \prec B$, то при любом M и при любом $\mathcal{P}_{B, M}$ -оцениваемом функционале ψ на M ядро B определяет $\mathcal{P}_{R, M}$ -н.л. оценку $\hat{\psi}_{B, M}$.*

Обратно, если $\mathcal{P}_{R, M}$ -н.л.о. $\hat{m}_{B, M}(t)$ определена для всех функционалов $\psi_i(m) = m(t)$ при любом одномерном линейном пространстве $M \subset H(B)$, то $R \prec B$.

Если $R \prec B$, $M \subset H(B)$ и при $P \in \mathcal{P}_{R, M}$ с P -вероятностью 1 $\xi(\cdot) \in H(B)$, то для любого функционала $\psi(m)$, $m \in M$, $\mathcal{P}_{B, M}$ -оцениваемого с помощью операций класса $LO(B)$, можно построить в $H(B)$ $\mathcal{P}_{R, M}$ -н.л. оценку наименьших квадратов (о.н.к.) $\hat{\psi}_{B, R} = \psi(\Pi_{\tilde{M}(B)}^{H(B)} \xi(\cdot))$. Пусть $g \in G_\psi(B)$ при $P \in \mathcal{P}_{R, M}$ с P -вероятностью 1

$$\psi_{B, R} = \langle \Pi_{\tilde{M}(B)}^{H(B)} \xi(\cdot), g(\cdot) \rangle_{H(B)} = \langle \xi(\cdot), \Pi_{\tilde{M}(B)}^{H(B)} g(\cdot) \rangle_{H(B)} = \hat{\psi}_{B, M}.$$

Пусть, например, $(T, \mathcal{B}, \mu)$ — пространство с σ -конечной мерой μ ; $\mathcal{E} = \{C: C \in \mathcal{B}, \mu(C) < \infty\}$; ядро $B_\mu(C, D) = \mu(C \cap D)$, $C, D \in \mathcal{E}$, можно рассматривать как «несобственное» воспроизводящее ядро пространства $L^2(T, \mathcal{B}, \mu)$: соответствие $f \leftrightarrow v$, где $v(C) = \int_C f d\mu$, $C \in \mathcal{E}$, есть изометрический изоморфизм между $L^2(T, \mathcal{B}, \mu)$ и $H(B_\mu)$. Условие $R \prec B_\mu$ равносильно тому, что $R(\cdot, t) \in L^2(T, \mathcal{B}, \mu)$ и $L_R^B: (L_R^B f)(t) = \int_T R(s, t) f(s) \mu(ds)$ — непрерывный оператор в $L^2(T, \mathcal{B}, \mu)$. Если $\xi(\cdot)$ — измеримая случайная функция на T и $\int_T E_P[\xi(t)]^2 \mu(dt) < \infty$, то $\xi(\cdot) \in L^2(T, \mathcal{B}, \mu)$ с P -вероятностью 1. О.н.к. в $L^2(T, \mathcal{B}, \mu)$ — о.н.к. с «весом» μ — определяются ядром β_μ .

Говорят, что $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}$ эффективна, если $\hat{\psi} = \hat{\psi}_{B, M}$; $\hat{\psi}$ состоит е л ь н а, если $D_P \psi = 0$ при всех $P \in \mathcal{P}_{R, M}$.

* Ср. с определением 3.1 в ⁽¹⁰⁾ (в формулировках условий 1) и 2) там вместо R следует читать $\tilde{R}$).

** Разумеется, $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}_{B, M}$ класса $LO(B)$ определяется также ядром B .

Теорема 3. Пусть $R < B$. Тогда: 1) для того чтобы $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}_{B, M}$ класса $LO(B)$ была эффективна (соответственно состоятельна), необходимо и достаточно, чтобы $L_R^B(\pi_\psi(\bar{B})) \subset \bar{M}(\bar{B}) \cap H(B)$ (соответственно $L_R^B(\pi_\psi(\bar{B})) = \emptyset$); 2) для того чтобы для всех функционалов ψ на M $\mathcal{P}_{R, M}$ -н.л.о. $\hat{\psi}_{B, M}$ класса $LO(B)$ были эффективны (соответственно состоятельны), необходимо и достаточно, чтобы $L_R^B(\bar{M}(B) \cap H(B)) \subseteq \bar{M}(B) \cap H(B)$ (соответственно $L_R^B(\bar{M}(B) \cap H(B)) = \{0\}$).

Пусть $T_\alpha \subset T$. Обозначим M^α семейство сужений $m^\alpha(\cdot)$ на T_α всех функций $m(\cdot)$ из M ; если $\psi(m_1) = \psi(m_2)$ коль скоро $m_1(t) = m_2(t)$ на T_α , мы можем рассмотреть сужение ψ^α функционала ψ на T_α : $\psi^\alpha(m^\alpha) = \psi(m)$. Легко убедиться, что всякая $\mathcal{P}_{R^\alpha, M^\alpha}$ -н.л.о. $\hat{\psi}^\alpha$ (по реализации $\xi(t)$, $t \in T_\alpha$) является $\mathcal{P}_{R, M}$ -н.л. оценкой функционала ψ .

Теорема 4. Пусть T_α , $\alpha \in A$, — возрастающая сеть множеств и $\bigcup_{\alpha \in A} T_\alpha = T$. Если $R < B$ и существуют $\mathcal{P}_{B^\alpha, M^\alpha}$ -оцениваемые сужения ψ^α функ-

ционала ψ на T_α при $\alpha \geq \alpha_0$, то при $\alpha > \alpha_0$ определены $\mathcal{P}_{B^\alpha, M^\alpha}$ -н.л.о. $\hat{\psi}_{B^\alpha, M^\alpha}^\alpha$

и при любой мере $P \in \mathcal{P}_{R, M}(P) \text{ l.i.m. } \hat{\psi}_{B^\alpha, M^\alpha}^\alpha = \hat{\psi}_{B, M}$ и следовательно,

$$\lim_{\alpha \in A} D_P \hat{\psi}_{B^\alpha, M^\alpha}^\alpha = D_P \hat{\psi}_{B, M}.$$

Условие $R < B$ необходимо для справедливости заключения теоремы 4 для функционалов $\psi_t(m) = m(t)$, $t \in T$, на всех одномерных подпространствах M в $H(B)$ (см. ^(4, 10)). Теорема 4 и критерий состоятельности, данный в теореме 3, обобщают теоремы о сходимости и состоятельности регрессионных оценок, принадлежащие Розанову (^{2, 3}) (см. также ⁽¹⁾), Холево (⁵) и автору (^{4, 10}). Некоторые следствия теоремы 4 (см. ⁽¹⁰⁾) близки к законам больших чисел для слабо зависимых случайных величин (см., например, ⁽⁸⁾, стр. 352) и к эргодическим теоремам для стационарных процессов и полей.

Институт физики и математики
Академии наук ЛитССР
Вильнюс

Поступило
26 II 1974

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ И. А. Ибрагимов, Ю. А. Розанов, Гауссовские случайные процессы, «Наука», 1970. ² Ю. А. Розанов, Гауссовские бесконечномерные распределения, «Наука», 1968. ³ Ю. А. Розанов, Советско-японский симпозиум по теории вероятностей, Хабаровск, август 1969, Новосибирск, Изд. АН СССР, 1969, стр. 231. ⁴ А. А. Темпельман, ДАН, т. 191, № 4, 772 (1970). ⁵ А. С. Холево, Теория вероятн. и ее примен., т. 14, 1, 73 (1969). ⁶ N. Aronszajn, Trans. Am. Math. Soc., v. 68, № 3, 337 (1957); Русск. пер. Математика, т. 7, № 2, 67 (1963). ⁷ E. Parzen, Proc. IV Berkeley Symp., v. 1, 1961, p. 479. ⁸ A. Renyi, Wahrscheinlichkeitsrechnung, Berlin, 1969. ⁹ Ju. A. Rozanov, Multivariate Analysis, v. 2, N. Y., 1969, p. 437. ¹⁰ A. A. Tempelman, The Second Symposium on Information Theory, Tsahkadzor, Armenia, USSR, Problems of Control and Information Theory, 1973, p. 329. ¹¹ N. D. Ylvisaker, Ann. Math. Statist., v. 33, 3, 1077 (1962).