Math-Net.Ru

Общероссийский математический портал

Ю. В. Малыхин, Условие выпуклости в теоремах
Кукера–Смейла в теории обучения, *Матем. заметки*,
2008, том 84, выпуск 1, 144–148

DOI: <https://doi.org/10.4213/mzm5198>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.82.47.9

26 апреля 2023 г., 16:30:09

Условие выпуклости в теоремах Кукера–Смейла в теории обучения

Ю. В. Малыхин

1. Введение. Пусть $X = \mathbb{R}^d$, $Y = \mathbb{R}$, ρ – неизвестная нам вероятностная мера на $Z = X \times Y$. Задача теории обучения, которую мы рассматриваем, состоит в том, чтобы по выборке $\mathbf{z} = \mathbf{z}^{(m)} = (z_1, \dots, z_m)$, состоящей из независимых векторов $z_i = (x_i, y_i)$, распределенных по мере ρ , построить функцию $f_{\mathbf{z}}: X \rightarrow Y$, минимизирующую величину ошибки

$$\mathcal{E}(f) := \int_{X \times Y} (f(x) - y)^2 d\rho.$$

Отображение $\mathbf{z} \mapsto f_{\mathbf{z}}$ называется *оценщиком*.

Пусть ρ_X – проекция меры ρ на X , т.е. $\rho_X(S) = \rho(S \times Y)$; $\rho(y|x)$ – условная мера. Минимум функционала $\mathcal{E}(\cdot)$ достигается на функции регрессии $f_\rho(x) = \int_Y y d\rho(y|x)$, как это следует из равенства

$$\mathcal{E}(f) - \mathcal{E}(f_\rho) = \|f - f_\rho\|_{L_2(\rho_X)}^2.$$

(Мы предполагаем, что f_ρ существует и лежит в $L_2(\rho_X)$.) Таким образом, задача сводится к аппроксимации функции f_ρ по метрике $L_2(\rho_X)$. Заметим, что, вообще говоря, мы не знаем ни f_ρ , ни ρ_X .

Обычно, исходя из знаний о мере ρ , выбирается класс $\mathcal{H}$, из которого будут браться функции $f_{\mathbf{z}}$. Минимизация функционала $\mathcal{E}(\cdot)$ на классе $\mathcal{H}$ равносильна нахождению функции из $\mathcal{H}$, ближайшей к f_ρ в метрике $L_2(\rho_X)$. Существование ближайшей функции обеспечивается требованием компактности $\mathcal{H}$; если ближайшая функция единственна, будем обозначать ее $f_{\mathcal{H}}$. В любом случае, корректно обозначение $\mathcal{E}(f_{\mathcal{H}}) := \inf_{f \in \mathcal{H}} \mathcal{E}(f)$.

Для построения оценщика естественно использовать метод наименьших квадратов (если минимум достигается на нескольких функциях, берем одну из них):

$$f_{\mathbf{z}} = f_{\mathbf{z}, \mathcal{H}} := \arg \min_{f \in \mathcal{H}} \frac{1}{m} \sum_{i=1}^m (f(x_i) - y_i)^2.$$

Нас будет интересовать вероятность того, что погрешность $f_{\mathbf{z}, \mathcal{H}}$ превзойдет минимальную погрешность больше, чем на заданное $\varepsilon > 0$:

$$\rho^m \{\mathbf{z} : \mathcal{E}(f_{\mathbf{z}, \mathcal{H}}) - \mathcal{E}(f_{\mathcal{H}}) > \varepsilon\}. \quad (1)$$

2. Результаты. В работе Кукера и Смейла [1] получены оценки сверху вероятностей (1). Там замечено, что условие выпуклости класса $\mathcal{H}$ позволяет доказывать более точные неравенства.

Через $N_\varepsilon(K, S)$ обозначается число элементов в минимальной ε -сети для множества K в метрическом пространстве S .

ТЕОРЕМА С [1]. Пусть $\mathcal{H}$ – компакт в $C(X)$. Предположим, что для всех $f \in \mathcal{H}$ выполнено $|f(x) - y| \leq M$ ρ -почти наверное. Положим

$$\sigma^2 := \sup_{f \in \mathcal{H}} \mathbb{D}((f(x) - y)^2).$$

Тогда для любого $\varepsilon > 0$ верно неравенство

$$\rho^m \{\mathbf{z} : \mathcal{E}(f_{\mathbf{z}, \mathcal{H}}) - \mathcal{E}(f_{\mathcal{H}}) > \varepsilon\} < 2N_{\varepsilon/(16M)}(\mathcal{H}, C(X)) \cdot \exp\left(-\frac{m\varepsilon^2}{8(4\sigma^2 + M^2\varepsilon/3)}\right).$$

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 08-01-00208).

ТЕОРЕМА С* [1]. Если выполнены условия теоремы С и класс $\mathcal{H}$ – выпуклый, то

$$\rho^m \{ \mathbf{z} : \mathcal{E}(f_{\mathbf{z}, \mathcal{H}}) - \mathcal{E}(f_{\mathcal{H}}) > \varepsilon \} < 2N_{\varepsilon/(24M)}(\mathcal{H}, C(X)) \cdot \exp\left(-\frac{m\varepsilon}{288M^2}\right). \quad (2)$$

Мы покажем, что в невыпуклом случае нельзя написать оценку со множителем вида $\exp(-m\varepsilon)$ вместо $\exp(-m\varepsilon^2)$. Это происходит из-за того, что у невыпуклого компакта $\mathcal{H}$ нарушается свойство единственности: имеются функции с более чем одной ближайшей в $\mathcal{H}$ (в метрике $L_2(\rho_X)$). Если же потребовать единственность в достаточно большой окрестности f_ρ , то неравенство, аналогичное (2), будет верно.

ТЕОРЕМА 1. Пусть μ – мера на X , $\mathcal{H}$ – невыпуклый компакт в $L_2(\mu)$ такой, что $|f| \leq M$ для всех $f \in \mathcal{H}$. Тогда существуют $M' \geq M$, $\gamma > 0$ такие, что для любого оценщика $\mathbf{z} \mapsto f_{\mathbf{z}} \in \mathcal{H}$, любого $\varepsilon \in (0, \varepsilon_0)$ и $m = \lfloor c/\varepsilon^2 \rfloor$ существует мера ρ , для которой $|y| \leq M'$ ρ -п.н., $\rho_X = \mu$ и

$$\rho^m \{ \mathbf{z} : \mathcal{E}(f_{\mathbf{z}}) - \mathcal{E}(f_{\mathcal{H}}) > \gamma \cdot \varepsilon \} > \frac{1}{14}.$$

Числа ε_0 и c суть абсолютные положительные константы.

ТЕОРЕМА 2. Пусть мера ρ и класс $\mathcal{H}$ таковы, что $\mathcal{H}$ компактен в $L_2(\rho_X)$ и в $C(X)$, $|f(x) - y| \leq M$ ρ -п.н. при $f \in \mathcal{H}$, и для некоторого $\alpha \in (0, 1)$ все функции из шара

$$\{f : \mathcal{E}(f) - \mathcal{E}(f_\rho) \leq \alpha^2 \cdot (\mathcal{E}(f_{\mathcal{H}}) - \mathcal{E}(f_\rho))\}$$

имеют единственную ближайшую из $\mathcal{H}$ по метрике $L_2(\rho_X)$. Тогда для любого $\varepsilon > 0$

$$\rho^m \{ \mathbf{z} : \mathcal{E}(f_{\mathbf{z}, \mathcal{H}}) - \mathcal{E}(f_{\mathcal{H}}) > \varepsilon \} < 2N_{\varepsilon/(24M)}(\mathcal{H}, C(X)) \cdot \exp\left(-\frac{\alpha m\varepsilon}{CM^2}\right),$$

где $C > 0$ абсолютная константа.

3. Доказательства теорем.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1. Всюду в доказательстве под $\|\cdot\|$ понимается норма пространства $L_2(\mu)$.

Частным случаем теоремы 2 из работы Царькова [2] является

УТВЕРЖДЕНИЕ 1. Пусть $\mathcal{H}$ – невыпуклый компакт в гильбертовом пространстве. Тогда в любом выпуклом множестве L , всюду плотном в этом пространстве, найдется точка $g \in L$, у которой по крайней мере два ближайших элемента в $\mathcal{H}$.

Применим это утверждение для гильбертова пространства $L_2(\mu)$, компакта $\mathcal{H}$ и множества $L = L_\infty(X)$. Получим функцию $g \in L_\infty(X)$ с двумя ближайшими $f_1, f_2 \in \mathcal{H}$. Введем обозначения

$$r := \|f_1 - f_2\| > 0, \quad R := \|f_1 - g\| = \|f_2 - g\| = \inf_{f \in \mathcal{H}} \|f - g\|.$$

Возьмем ε из формулировки теоремы. Определим меру ρ_1 условиями $(\rho_1)_X = \mu$ и

$$\rho_1(y|x) = \begin{cases} \frac{1}{2} + \frac{\varepsilon}{2}, & y = g(x) + (f_1(x) - f_2(x)), \\ \frac{1}{2} - \frac{\varepsilon}{2}, & y = g(x) - (f_1(x) - f_2(x)). \end{cases}$$

Аналогично определим ρ_2 , поменяв местами f_1 и f_2 . Заметим, что $f_{\rho_1} = g + \varepsilon(f_1 - f_2)$ и $f_{\rho_2} = g - \varepsilon(f_1 - f_2)$.

По построению $(\rho_i)_X = \mu$ и $|y| \leq M' := \|g\|_\infty + 2M$. Мы докажем, что неравенство теоремы выполнено с $\gamma = r^2/4$ для хотя бы одной из мер ρ_i , $i = 1, 2$. Нас интересуют множества

$$U_i := \{f \in \mathcal{H} : \mathcal{E}_i(f) - \mathcal{E}_i(f_{\mathcal{H}}) \leq \gamma \cdot \varepsilon\}, \quad i = 1, 2,$$

где $\mathcal{E}_i(\cdot)$ и $\mathcal{E}_i(f_{\mathcal{H}})$ обозначают, соответственно, $\mathcal{E}(\cdot)$ и $\mathcal{E}(f_{\mathcal{H}})$ для меры $\rho = \rho_i$. Можно избавиться от $\mathcal{E}_i(\cdot)$, обозначив $d_i := \inf_{f \in \mathcal{H}} \|f_{\rho_i} - f\|$ и записав

$$U_i := \{f \in \mathcal{H} : \|f - f_{\rho_i}\|^2 - d_i^2 \leq \gamma \cdot \varepsilon\}.$$

(На всякий случай напомним, что $\|\cdot\|$ обозначает норму пространств $L_2(\mu) = L_2((\rho_i)_X)$, $i = 1, 2$.) Покажем, что $U_1 \cap U_2 = \emptyset$ при $\varepsilon \in (0, 1/4)$ и $\gamma = r^2/4$. По определению

$$d_i \leq \|f_{\rho_i} - f_i\| = \sqrt{R^2 - \varepsilon r^2 + \varepsilon^2 r^2}.$$

С другой стороны, $d_i \geq R - \varepsilon r > R - r/4 \geq R/2$. Для любого $f \in U_i$

$$\begin{aligned} (r^2/4)\varepsilon &\geq \|f - f_{\rho_i}\|^2 - d_i^2 \\ &= (\|f - f_{\rho_i}\| + d_i)(\|f - f_{\rho_i}\| - d_i) \geq R \cdot (\|f - f_{\rho_i}\| - d_i), \\ \|f - f_{\rho_i}\| &\leq \frac{r^2}{4R}\varepsilon + d_i \leq \frac{r^2}{4R}\varepsilon + \sqrt{R^2 - \varepsilon r^2 + \varepsilon^2 r^2} < R. \end{aligned}$$

В последнем неравенстве использовано, что $\varepsilon < 1/2$. Если $f \in U_1 \cap U_2$, то $\|f - f_{\rho_1}\| < R$, $\|f - f_{\rho_2}\| < R$, откуда $\|f - g\| < R$, что невозможно.

Нам нужно показать, что $f_{\mathbf{z}} \notin U_i$ с вероятностью $\rho_i^m > 1/14$ хотя бы для одного $i \in \{1, 2\}$. Предположим, это неверно. Поскольку $U_1 \cap U_2 = \emptyset$, то

$$\begin{aligned} P_1 &= \rho_1^m \{\mathbf{z} : f_{\mathbf{z}} \in U_1\} \geq \frac{13}{14}, \\ P_2 &= \rho_2^m \{\mathbf{z} : f_{\mathbf{z}} \in U_1\} \leq \frac{1}{14}. \end{aligned}$$

Выведем из последних двух неравенств противоречие.

Опишем, как получить случайный вектор $\mathbf{z} \in Z^m$ с распределением ρ_1^m . Сначала m раз подкинем $(1/2 + \varepsilon/2)$ -монетку (вероятность орла равна $1/2 + \varepsilon/2$). Пусть выпало s орлов. Далее возьмем m независимых случайных точек $x_1, \dots, x_m \in X$ с распределением μ . Для s индексов, соответствующих орлам, положим $y_i = g(x_i) + (f_1(x_i) - f_2(x_i))$, а для остальных индексов $y_i = g(x_i) - (f_1(x_i) - f_2(x_i))$. Распределение полученного вектора $\mathbf{z} = ((x_1, y_1), \dots, (x_m, y_m))$ и будет ρ_1^m . (Иными словами, $y_i = g(x_i) + \xi_i(f_1(x_i) - f_2(x_i))$, где $x_1, \dots, x_m \in X$ и $\xi_1, \dots, \xi_m \in \{-1, 1\}$ независимы, распределение x_j есть μ и $P(\xi_j = \pm 1) = 1/2 \pm \varepsilon$.)

Можно сделать по другому: сгенерируем случайную величину s , распределенную по биномиальному закону $\mathcal{B}(m, 1/2 + \varepsilon/2)$, возьмем, как раньше, m случайных точек x_i и выберем s случайных индексов. Для выбранных индексов $y_i = g(x_i) + (f_1(x_i) - f_2(x_i))$, для остальных $y_i = g(x_i) - (f_1(x_i) - f_2(x_i))$.

В построении случайного вектора $\mathbf{z} \in Z^m$ с распределением ρ_2^m конструкция та же, только биномиальный закон $\mathcal{B}(m, 1/2 - \varepsilon/2)$. Отметим, что после того, как s выбрано, построение для ρ_1 и ρ_2 одинаково. В силу этого корректно обозначение

$$q(s_0) = \rho_i^m \{\mathbf{z} : f_{\mathbf{z}} \in U_1 \mid s = s_0\}$$

и вероятности P_1, P_2 можно записать в виде

$$P_i = \sum_{s=0}^m p_i(s) q(s),$$

где

$$p_1(s) = \binom{m}{s} \left(\frac{1}{2} + \frac{\varepsilon}{2}\right)^s \left(\frac{1}{2} - \frac{\varepsilon}{2}\right)^{m-s}, \quad p_2(s) = p_1(m-s).$$

Рассмотрим участок $s \in [m(1/2 - \varepsilon/2), m(1/2 + \varepsilon/2)]$. На нем $p_1(s)$ возрастает с ростом s , а $p_2(s)$ убывает. Если мы переупорядочим $q(s)$ на этом участке в порядке возрастания, то $P_1 = \sum p_1(s)q(s)$ не уменьшится, $P_2 = \sum p_2(s)q(s)$ не увеличится; будет по-прежнему $P_1 \geq 13/14$, $P_2 \leq 1/14$.

Если ξ_n – случайная величина, распределенная по биномиальному закону $\mathcal{B}(n, 1/2 + \varepsilon/2)$, и $n = n(\varepsilon) = \lfloor c/\varepsilon^2 \rfloor$, то в силу центральной предельной теоремы

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P} \left\{ n \left(\frac{1}{2} - \frac{\varepsilon}{2} \right) \leq \xi_n \leq \frac{n}{2} \right\} = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{c}}^{2\sqrt{c}} e^{-t^2/2} dt.$$

Можно подобрать c таким, чтобы выражение в правой части было больше $1/7$.

Обозначим $\Delta_1 := [m(1/2 - \varepsilon/2), m/2]$, $\Delta_2 := [m/2, m(1/2 + \varepsilon/2)]$. В силу вышесказанного и соотношений $m = \lfloor c/\varepsilon^2 \rfloor$, $\varepsilon < \varepsilon_0$ имеем

$$\sum_{s \in \Delta_1} p_1(s) > \frac{1}{7}, \quad \sum_{s \in \Delta_2} p_2(s) > \frac{1}{7}.$$

Если $q(\lfloor m/2 \rfloor) > 1/2$, то $P_2 \geq \sum_{s \in \Delta_2} p_2(s)q(s) > 1/14$, что невозможно. Если же $q(\lfloor m/2 \rfloor) \leq 1/2$, то $a := \sum_{s \in \Delta_1} p_1(s)q(s) > 1/7$,

$$P_1 = \sum_{s \in \Delta_1} p_1(s)q(s) + \sum_{s \notin \Delta_1} p_1(s)q(s) \leq \frac{a}{2} + (1-a) < \frac{13}{14}.$$

Мы пришли к противоречию и, тем самым, доказали теорему.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 2 повторяет доказательство теоремы C* за исключением того, что вместо неравенства

$$\|f - f_\rho\|_{L_2(\rho_X)}^2 \leq \mathcal{E}(f) - \mathcal{E}(f_{\mathcal{H}}), \quad f \in \mathcal{H},$$

обеспеченного выпуклостью $\mathcal{H}$ (лемма 5 [1]), используется неравенство

$$\|f - f_\rho\|_{L_2(\rho_X)}^2 \leq \left(1 + \frac{1}{\alpha}\right) \cdot (\mathcal{E}(f) - \mathcal{E}(f_{\mathcal{H}})), \quad f \in \mathcal{H}. \quad (3)$$

Докажем (3). Множество функций $f \in L_2(\rho_X)$, не удовлетворяющих этому неравенству, образует открытый шар радиуса $(1 + \alpha) \cdot \|f_\rho - f_{\mathcal{H}}\|$ с центром в $f_\rho + \alpha(f_\rho - f_{\mathcal{H}})$. В этом можно убедиться, проведя плоскость через f , f_ρ и $f_{\mathcal{H}}$ и записав неравенство в координатах. В указанном шаре нет функций из $\mathcal{H}$, как показывает следующее утверждение (см., например, [3; §2]).

УТВЕРЖДЕНИЕ 2. Пусть $\mathcal{H}$ – компакт в гильбертовом пространстве V , $x \in V \setminus \mathcal{H}$, $x' \in \mathcal{H}$ – ближайший элемент в $\mathcal{H}$ к x , $R = \|x - x'\|$. Если в шаре $\{y \in V : \|x - y\| \leq \alpha R\}$, $\alpha \in (0, 1)$, все точки обладают единственным ближайшим в $\mathcal{H}$, то x' является ближайшим к $x + \alpha(x - x')$.

ЗАМЕЧАНИЕ. В теореме 2 можно требовать единственность не во всем шаре, а в пересечении его с $L_\infty(\rho_X)$. Это следует из того, что в утверждении 2 достаточно требовать единственность в пересечении шара с произвольным всюду плотным выпуклым множеством (Царьков, устное сообщение; при доказательстве используется техника из [2], [4]).

Автор благодарит С.В. Конягина за постановку задач и внимание к работе, а также И.Г. Царькова за обсуждение необходимых результатов из геометрической теории приближений.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

- [1] F. Cucker, S. Smale, *Bull. Amer. Math. Soc. (N.S.)*, **39**:1 (2002), 1–49. [2] И. Г. Царьков, *Матем. заметки*, **75**:2 (2004), 287–301. [3] Л. П. Власов, *УМН*, **28**:6 (1973), 3–66. [4] И. Г. Царьков, *Матем. заметки*, **48**:4 (1990), 122–131.

Ю. В. Малыхин

Математический институт им. В. А. Стеклова РАН

Поступило

10.11.2007