ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—март 2017. Том 24, № 1. С. 31–55

УДК 519.715 DOI: 10.17377/daio.2017.24.508

О РАЗРЕШАЮЩЕМ МНОЖЕСТВЕ 2-ПОРОГОВОЙ ФУНКЦИИ ДВУХ ПЕРЕМЕННЫХ

Е. М. Замараева

Нижегородский гос. университет им. Н. И. Лобачевского,
пр. Гагарина, 23, 603950 Нижний Новгород, Россия
E-mail: elena.zamaraeva@gmail.com

Аннотация. Рассматриваются k-пороговые функции n переменных, т. е. функции, представимые в виде конъюнкции k пороговых функций. Для случая $n = 2$, $k = 2$ даются верхние оценки мощности тупикового разрешающего множества функции в зависимости от её различных свойств. Ил. 6, библиогр. 9.

Ключевые слова: машинное обучение, пороговая функция, длина обучения, разрешающее множество.

1. Введение

Пусть $E_n^d = \{0, 1, \ldots, n-1\}^d$, $n \geqslant 2$, $d \geqslant 1$ и k — натуральное число. Функция $f\colon E_n^d \to \{0, 1\}$ называется *k-пороговой*, если существуют действительные числа $a_{10}, a_{11}, \ldots, a_{kd}$ такие, что

$$M_1(f) = \left\{x \in E_n^d \mid \sum_{j=1}^{d} a_{ij}x_j \leqslant a_{i0},\ i = 1, \ldots, k\right\},$$

где $M_\nu(f) = \{x \mid f(x) = \nu\}$, $\nu = 0, 1$. Неравенства $\sum\limits_{j=1}^{d} a_{ij}x_j \leqslant a_{i0}$, $i = 1, \ldots, k$, называются *пороговыми*, или *определяющими* функцию f. 1-Пороговая функция называется просто *пороговой*.

Обозначим через $\mathfrak{T}(d, n, k)$ и $\mathfrak{T}(d, n)$ классы k-пороговых и пороговых функций, определённых на E_n^d, соответственно. Заметим, что при $j > k$ имеет место $\mathfrak{T}(d, n, k) \subseteq \mathfrak{T}(d, n, j)$. Обозначим $\mathfrak{T}(d, n, *) = \bigcup\limits_{k \geqslant 1} \mathfrak{T}(d, n, k)$.

Для любой k-пороговой функции f существует (вообще говоря, не единственное) множество пороговых функций $\{f_1, \ldots, f_k\}$ такое, что $f(x) = f_1(x) \& \ldots \& f_k(x)$. Будем говорить, что f *определяется* функциями $f_1, \ldots, f_k$ и $\{f_1, \ldots, f_k\}$ — *определяющее множество* f.

Через $\mathrm{Conv}(X)$ обозначим выпуклую оболочку множества точек $X \subseteq \mathbb{R}^d$. Для функции $f\colon E_n^d \to \{0,1\}$ обозначим $P(f) = \mathrm{Conv}(M_1(f))$. Для любого выпуклого многогранника P с вершинами из E_n^d существует единственная функция $f \in \mathfrak{T}(d,n,*)$ такая, что $P = P(f)$. Более того, существует взаимно однозначное соответствие между функциями из класса $\mathfrak{T}(d,n,*)$ и выпуклыми многогранниками с вершинами из E_n^d. Множество вершин выпуклого многогранника P обозначим через $\mathrm{Vert}(P)$.

Для обозначения асимптотических верхних и точных оценок скорости роста функций будем использовать стандартные обозначения O и Θ соответственно. Формально для функций $g(n)$ и $h(n)$ запись $g(n) = O(h(n))$ означает, что найдутся константа $c > 0$ и n_0 такие, что $|g(n)| \leqslant c|h(n)|$ для всех $n > n_0$. Аналогично запись $g(n) = \Theta(h(n))$ означает, что найдутся такие $c_1, c_2 > 0$ и n_0, что $c_1 h(n) \leqslant g(n) \leqslant c_2 h(n)$.

Пусть $\mathcal{C}$ — класс $\{0,1\}$-значных функций над E_n^d и $f \in \mathcal{C}$. *Разрешающим множеством* функции f относительно класса $\mathcal{C}$ называется множество $T \subseteq E_n^d$ такое, что никакая другая функция из $\mathcal{C}$ не совпадает с f на всём T. Разрешающее множество T называется *тупиковым*, если никакое его собственное подмножество не является разрешающим для f.

Обозначим через $\sigma(f,\mathcal{C})$ и назовём *длиной обучения функции f относительно класса $\mathcal{C}$* минимальную мощность разрешающего множества f относительно $\mathcal{C}$. *Длиной обучения класса $\mathcal{C}$* назовём

$$\sigma(\mathcal{C}) = \max_{f\in\mathcal{C}} \sigma(f,\mathcal{C}).$$

Точка $x \in E_n^d$ называется *существенной* точкой функции $f \in \mathcal{C}$ относительно класса $\mathcal{C}$, если существует функция $g \in \mathcal{C}$, совпадающая с f всюду в E_n^d кроме точки x. Обозначим множество существенных точек f относительно класса $\mathcal{C}$ через $S(f,\mathcal{C})$ или $S(f)$, если $\mathcal{C}$ ясен из контекста. Пусть $S_\nu(f) = S(f) \cap M_\nu(f)$.

Известно (см., например, [4, 7]), что множество существенных точек пороговой функции является разрешающим. Кроме того, из [1, 2, 6, 7] следует, что при любом фиксированном $d \geqslant 2$

$$\sigma(\mathfrak{T}(d,n)) = \Theta\bigl(\log_2^{d-2} n\bigr),$$

при $d = 2$ мощность тупикового разрешающего множества пороговой функции равна 3 или 4, а средняя мощность этого множества асимптотически стремится к $\frac{7}{2}$ [3].

В [9] рассмотрены комбинаторные и структурные свойства разрешающих множеств k-пороговых функций для $k \geqslant 2$. В частности, описан

подкласс функций $\{f \in \mathfrak{T}(d,n,*) \mid |M_1(f)| \leqslant 1\}$, для которых

$$\sigma(f, \mathfrak{T}(d,n,*)) = \Theta(n^d). \tag{1}$$

Также для функций $f \in \mathfrak{T}(2,n,*)$ с ненулевой площадью $P(f)$ найдена неулучшаемая верхняя оценка длины обучения

$$\sigma(f, \mathfrak{T}(2,n,*)) = O(n). \tag{2}$$

В [5] изучался класс k-HALFSPACES$^2_{n,p} \subset \mathfrak{T}(2,n,k)$, где $0 < p \leqslant \frac{\pi}{2}$. Класс k-HALFSPACES$^2_{n,p}$ состоит из функций $f \in \mathfrak{T}(2,n,k)$, для которых многоугольник $P(f)$ удовлетворяет двум дополнительным условиям: во-первых, каждое ребро $P(f)$ имеет длину не меньше $16\lceil\frac{1}{p}\rceil$, во-вторых, угол α между любой парой смежных рёбер удовлетворяет неравенству $p \leqslant \alpha \leqslant \pi - p$. Из [5] следует, что

$$\sigma\bigl(k\text{-HALFSPACES}^2_{n,p}\bigr) = O\left(k\left(\frac{1}{p} + \log n\right)\right). \tag{3}$$

В данной работе рассматривается класс $\mathfrak{T}(2,n,2)$. Приводится оценка длины обучения функций, аналогичная (3), без накладываемых в [5] ограничений на функции, хотя некоторые специальные случаи функций исключаются. Кроме того, оценка (2) расширяется до всех 2-пороговых функций f, для которых $|M_1(f)| > 1$.

В разд. 2 приводятся основные результаты из [9], касающиеся класса $\mathfrak{T}(2,n,2)$ с некоторыми модификациями, а также несколько вспомогательных геометрических утверждений, которые используются позднее. Разд. 3 посвящён верхней оценке длины обучения функций из класса $\mathfrak{T}(2,n,*)$ в зависимости от n. Основной результат излагается в разд. 4, где приводится верхняя оценка длины обучения 2-пороговой функции в зависимости от её геометрических свойств.

2. Предварительные результаты

Приведём несколько вспомогательных утверждений, касающихся выпуклых объектов на плоскости, которые понадобятся нам в дальнейшем. Тривиальные утверждения приведём без доказательств.

Если a — множество точек, образующее кривую или объединение непересекающихся кривых, то через $l(a)$ обозначим суммарную длину всех кривых из a. Обозначим через $\mathcal{P}(P)$ и $\mathcal{S}(P)$ периметр и площадь выпуклого многоугольника P соответственно, а через $b(P)$ — границу P.

Утверждение 1. *Пусть* A, B, C *— целые точки, не лежащие на одной прямой, такие, что* $\angle ACB \leqslant \angle BAC \leqslant \angle ABC$. *Тогда* $\sin \angle ACB \geqslant \frac{1}{l(AC)^2}$.

Утверждение 2. *Пусть* $\triangle ABC$ *— треугольник, не имеющий внутренних целых точек, и* $\angle ACB = \angle ABC = \alpha$. *Тогда* $l(BC) \leqslant \sqrt{2}\operatorname{ctg}\frac{\alpha}{2}$.

Утверждение 3. *Для произвольного* $\triangle ABC$ *и любой точки* $D \in \triangle ABC$ *верно неравенство* $\angle ABC \leqslant \angle ADC$.

Также в ходе дальнейших рассуждений понадобятся следующие тригонометрические неравенства, справедливые при $0 < x < \frac{\pi}{2}$:

$$\operatorname{ctg} x < \frac{1}{x}, \tag{4}$$

$$\sin x > \frac{2x}{\pi}. \tag{5}$$

Пусть P — выпуклый многоугольник на плоскости (отрезок в вырожденном случае), и пусть $a_1x_1 + a_2x_2 = a_0$ — уравнение прямой, на которой лежит его ребро e. Назовём *рёберным неравенством* ребра e то из неравенств

$$a_1x_1 + a_2x_2 \leqslant a_0, \quad a_1x_1 + a_2x_2 \geqslant a_0,$$

которое удовлетворяется для всех точек P. Заметим, что если P — отрезок, то он имеет единственное ребро и любое нестрогое неравенство, обращающееся в равенство на P, является его рёберным неравенством.

Пусть $a_1x_1 + a_2x_2 \leqslant a_0$ — рёберное неравенство ребра e многоугольника P. Рассмотрим ближайшую параллельную e прямую, содержащую хотя бы одну целую точку и такую, что указанное рёберное неравенство в ней не выполняется. Нестрогое неравенство, обращающееся в равенство на этой прямой и выполняющееся во всех точках P, назовём *расширенным рёберным неравенством* ребра e.

Расширением P назовём множество решений системы:

$$P' = \{x = (x_1, x_2) \mid a'_{i1}x_1 + a'_{i2}x_2 \leqslant a'_{i0},\ i = 1, \ldots, |\mathrm{Vert}(P)|\},$$

где $a'_{i1}x_1 + a'_{i2}x_2 \leqslant a'_{i0},\ i = 1, \ldots, |\mathrm{Vert}(P)|$, — система расширенных рёберных неравенств для рёбер P максимальной мощности без неравенств следствий.

Для произвольного выпуклого многоугольника P и его вершины v через $q(v, P)$ обозначим угол при вершине v в этом многоугольнике.

Утверждение 4. *Пусть P — выпуклый многоугольник с целочисленными вершинами и P' — его расширение. Тогда*

$$\mathcal{P}(P') < 5\mathcal{P}(P) + \frac{4}{\sin \min\limits_{v \in \mathrm{Vert}(P)} \frac{q(v,P)}{2}}.$$

Доказательство. Рассмотрим множество всех точек таких, что если какое-то рёберное неравенство P в точке не выполняется, то она находится на расстоянии не больше 1 от прямой, на которой лежит соответствующее ребро. Это множество точек образует выпуклый многоугольник P'' (рис. 1). Заметим, что все точки P' также удовлетворяют указанному условию, значит, $P' \subseteq P''$, следовательно,

$$\mathcal{P}(P') \leqslant \mathcal{P}(P''). \tag{6}$$

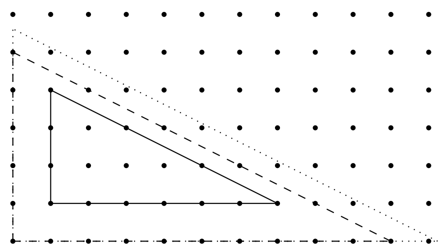

Рис. 1. P (сплошная линия), P' (штриховая линия) и P'' (пунктирная линия)

Используя определение P'', нетрудно заключить, что

$$\mathcal{P}(P'') = \mathcal{P}(P) + 2 \sum_{v \in \mathrm{Vert}(P)} \operatorname{ctg} \frac{q(v,P)}{2}. \tag{7}$$

Так как в выпуклом многоугольнике не больше двух внутренних углов может быть меньше $\frac{\pi}{3}$ и $|\mathrm{Vert}(P)| \leqslant \mathcal{P}(P)$, можно преобразовать последнее слагаемое:

$$2 \sum_{v \in \mathrm{Vert}(P)} \operatorname{ctg} \frac{q(v,P)}{2} \leqslant 2\left((\mathcal{P}(P) - 2) \operatorname{ctg} \frac{\pi}{6} + 2 \operatorname{ctg} \min_{v \in \mathrm{Vert(P)}} \frac{q(v,P)}{2} \right)$$
$$\leqslant 2\sqrt{3}(\mathcal{P}(P) - 2) + \frac{4}{\sin \min\limits_{v \in \mathrm{Vert(P)}} \frac{q(v,P)}{2}}. \tag{8}$$

В последнем неравенстве пользуемся тем фактом, что $\operatorname{ctg} x \leqslant \frac{1}{\sin x}$ для любого $x \in (0, \pi)$. Требуемое следует из (6)–(8). Утверждение 4 доказано.

Следствие 1. *Пусть выпуклый многоугольник P с целочисленными вершинами вписан в некоторый квадрат E со сторонами, параллельными осям координат, а P' — его расширение. Тогда*

$$l(b(P' \setminus E)) < 5\mathcal{P}(P) + \frac{4}{\sin \min\limits_{v \in \mathrm{Vert}(P) \setminus b(E)} \frac{q(v,P)}{2}} + 8.$$

Для заданного выпуклого многоугольника P назовём *опорной* прямую, проходящую через хотя бы одну из вершин P таким образом, что P лежит целиком в одной (замкнутой) полуплоскости, ограниченной этой прямой. Обозначим через $\mathrm{Diam}(P)$ и назовём *диаметром* P максимальное расстояние между двумя параллельными опорными прямыми P. Обозначим через $w(P)$ и назовём *шириной* P минимальное расстояние между двумя параллельными опорными прямыми P.

Утверждение 5. *Пусть P — выпуклый многоугольник с целочисленными вершинами и P' — его расширение. Тогда расстояние от прямой, проходящей через наиболее удалённые друг от друга точки P, до любой точки P' не превосходит $w(P) + 2$.*

Доказательство. Пусть $v_1, v_2 \in \mathrm{Vert}(P)$ такие, что

$$l(v_1v_2) = \mathrm{Diam}(P).$$

Отрезок v_1v_2 делит P на два выпуклых многоугольника, обозначим их через P_1 и P_2. Заметим, что

$$\mathrm{Diam}(P) = \mathrm{Diam}(P_1) = \mathrm{Diam}(P_2), \quad w(P) \geqslant \max(w(P_1), w(P_2)).$$

Так как $l(v_1v_2) = \mathrm{Diam}(P_1)$, имеем $\angle v_1vv_2 \geqslant \max(\angle v_1v_2v, \angle v_2v_1v)$ для любой $v \in \mathrm{Vert}(P_1) \setminus \{v_1, v_2\}$, а значит, $q(v, P) = q(v, P_1) \geqslant \frac{\pi}{3}$. Кроме того, высота, опущенная из v на основание v_1v_2, является шириной треугольника $\triangle vv_1v_2$ и не может превосходить $w(P_1)$. Отсюда следует, что все точки P_1 находятся на расстоянии не более $w(P)$ от прямой, проходящей через v_1v_2. Поскольку аналогичные рассуждения справедливы и для P_2, все точки многоугольника P находятся на расстоянии не больше $w(P)$ от прямой, проходящей через v_1v_2, и все углы, кроме, быть может, углов при v_1 и v_2, не меньше $\frac{\pi}{3}$. Если построить многоугольник P'' так же, как это делается в утверждении 4, то нетрудно увидеть, что все точки P'', а значит, и все точки P', находятся на расстоянии не больше $w(P) + 2$ от прямой, проходящей через v_1v_2. Утверждение 5 доказано.

Введём обозначение множества точек E_n^2, лежащих на границе его выпуклой оболочки:

$$B\big(E_n^2\big) = \big\{x = (x_1, x_2) \in E_n^2 \mid x_1 = 0 \vee x_2 = 0 \vee x_1 = n-1 \vee x_2 = n-1\big\}.$$

Пусть $f \in \mathfrak{T}(2, n, *)$ и $|M_1(f)| > 1$. Через $P'(f)$ обозначим расширение $P(f)$ и введём обозначение

$$\Delta P(f) = P'(f) \setminus P(f).$$

Для полноты изложения приведём следующие результаты [9]. Точку x в пространстве $\mathbb{Z}^d$ назовём *видимой* из d-мерного политопа P, если выполнено равенство

$$(\mathrm{Conv}(P \cup \{x\}) \setminus P) \cap \mathbb{Z}^d = \{x\}.$$

Теорема 1 [9]. *Пусть $f \in \mathfrak{T}(d, n, *)$, $d \geqslant 2$, $n \geqslant 2$. Тогда множество*

$$S(f) = \begin{cases} E_n^d, & f \equiv 0,\\ \mathrm{Vert}(P(f)) \cup \{x \mid \mathrm{Conv}(P(f) \cup \{x\}) \cap M_0(f) = \{x\}\}, & f \not\equiv 0, \end{cases}$$

является единственным тупиковым разрешающим множеством функции f относительно $\mathfrak{T}(d, n, *)$.

Утверждение 6 [9]. *Пусть $f \in \mathfrak{T}(2, n, *)$ и $\mathcal{S}(P(f)) > 0$. Тогда*

$$\{x \mid \mathrm{Conv}(P(f) \cup \{x\}) \cap M_0(f) = \{x\}\} = \Delta P(f) \cap M_0(f).$$

Из теоремы 1 следует, что множество, состоящее из вершин $P(f)$ и всех видимых из $P(f)$ точек E_n^d, является разрешающим множеством ненулевой k-пороговой функции при любых k. Утверждение 6 уточняет, что при $d = 2$ множество видимых из $P(f)$ точек совпадает с множеством целых точек из $\Delta P(f)$.

Введём обозначение $q_{\min}(P(f)) = \min\limits_{v \in \mathrm{Vert}(P(f)) \setminus B(E_n^2)} q(v, P(f))$.

Лемма 1. *Пусть $f \in \mathfrak{T}(2, n, *)$ и $\mathcal{S}(P(f)) > 0$. Тогда*

$$\begin{aligned} S(f) &= \mathrm{Vert}(P(f)) \cup (\Delta P(f) \cap M_0(f)),\\ |S(f)| &= O\bigg(\mathcal{P}(P(f)) + \frac{1}{q_{\min}(P(f))}\bigg). \end{aligned} \tag{9}$$

Доказательство. Из теоремы 1 и утверждения 6 следует, что

$$S(f) = \mathrm{Vert}(P(f)) \cup (\Delta P(f) \cap M_0(f)).$$

Заметим, что $S_1(f) = \mathrm{Vert}(P(f))$, а значит, $|S_1(f)| = |\mathrm{Vert}(P(f))| \leqslant \mathcal{P}(P(f))$. Далее $S_0(f) = b(P'(f)) \cap E_n^2$. Так как $b(P'(f)) \cap \mathrm{Conv}\big(E_n^2\big)$ представляет собой либо замкнутый контур, либо объединение не более четырёх непересекающихся кривых, количество целых точек в этом множестве не превосходит $l\big(b(P'(f)) \cap \mathrm{Conv}\big(E_n^2\big)\big) + 4$. Используя следствие 1, получаем

$$\begin{aligned} |S(f)| \leqslant \mathcal{P}(P(f)) + l\big(b(P'(f)) \cap \mathrm{Conv}\big(E_n^2\big)\big) + 4 \\ \leqslant 6\mathcal{P}(P(f)) + \frac{4}{\sin \frac{q_{\min}(P(f))}{2}} + 12, \end{aligned}$$

откуда с применением неравенства (5) следует утверждение леммы. Лемма 1 доказана.

Замечание 1. В [9] лемма 1 доказана для угла

$$q'_{\min}(P(f)) = \min_{v \in \mathrm{Vert}(P(f))} q(v, P(f)).$$

Так как $q'_{\min}(P(f)) \leqslant q_{\min}(P(f))$, лемма 1 является усилением результата из [9].

В дальнейшем понадобятся следующие утверждения из [9].

Лемма 2 [9]. *Пусть $f \in \mathfrak{T}(2, n, 2)$, $M_1(f) \cap B\big(E_n^2\big) \neq \varnothing$ и для некоторого множества $\{f_1, f_2\}$ пороговых функций, определяющих f, выполнено*

$$S(f_1) \cap M_0(f_2) = \varnothing, \quad S(f_2) \cap M_0(f_1) = \varnothing.$$

Тогда $\{f_1, f_2\}$ — единственное определяющее множество f.

Лемма 3 [9]. *Пусть $f \in \mathfrak{T}(2, n, 2)$ и $\{f_1, f_2\}$ — единственное определяющее множество f. Тогда $\sigma(f, \mathfrak{T}(2, n, 2)) \leqslant 9$, если $M_1(f) \cap B\big(E_n^2\big) \neq \varnothing$.*

Следствие 2 [9]. *Лемма 3 также справедлива, когда пространство является подмножеством E_n^2 и его выпуклая оболочка не содержит точек из E_n^2 вне этого пространства.*

3. Длина обучения в классе $\mathfrak{T}(2, n, *)$

В [9] оценивается рост длины обучения k-пороговой функции относительно класса $\mathfrak{T}(2, n, *)$ в зависимости от n. Оценка даётся для функций, принимающих значение 1 не более чем в одной точке, и для функций с ненулевой площадью $P(f)$. Таким образом, остались нерассмотренными лишь функции, для которых $P(f)$ является отрезком, и следующая лемма обращается именно к ним.

Лемма 4. *Пусть $f \in \mathfrak{T}(2, n, *)$, $|M_1(f)| > 1$ и существуют целые a, b, c такие, что* НОД$(a, b) = 1$ *и $ax_1 + bx_2 + c = 0$ для любых $l(x_1, x_2) \in M_1(f)$. Тогда*

$$S(f) = \{v_1, v_2\} \cup \{v_1 \pm (b, -a), v_2 \pm (b, -a)\} \cap M_0(f) \\ \cup \{(x_1, x_2) \in M_0(f) \mid |ax_1 + bx_2 + c| = 1\},$$

где $\mathrm{Vert}(P(f)) = \{v_1, v_2\}$, *и* $|S(f)| \leqslant \frac{2n}{\max(|a|,|b|)} + 4$.

Доказательство. Так как $|M_1(f)| > 1$ и все точки $M_1(f)$ лежат на одной прямой, $P(f)$ является отрезком и $|\mathrm{Vert}(P(f))| = 2$. По теореме 1 $S_1(f) = \mathrm{Vert}(P(f)) = \{v_1, v_2\}$ и $S_0(f)$ состоит из нулей функций, видимых из $P(f)$. Если рассмотреть все целочисленные точки на прямой $ax_1 + bx_2 + c = 0$, то только две из них могут войти в $S_0(f)$ — это ближайшие к отрезку $P(f)$ нули функции, расположенные на прямой $ax_1 + bx_2 + c = 0$, и множество этих точек можно представить в виде

$$\{v_1 \pm (b, -a), v_2 \pm (b, -a)\} \cap M_0(f). \tag{10}$$

Так как НОД$(a, b) = 1$, по формуле Пика (см., например, [8]) среди целых точек, не лежащих на прямой $ax_1+bx_2+c = 0$, все целые точки на прямых $ax_1 + bx_2 + c + 1 = 0$ и $ax_1 + bx_2 + c - 1 = 0$ и только они видимы из $P(f)$. Отсюда и из (10) получаем

$$S_0(f) = \{v_1 \pm (b, -a), v_2 \pm (b, -a)\} \cap M_0(f) \\ \cup \{(x_1, x_2) \in M_0(f) \mid |ax_1 + bx_2 + c| = 1\}.$$

В квадрат со стороной n и целочисленными вершинами на прямой $ax_1+bx_2+c = 0$ для взаимно простых a и b помещается не более $\frac{n}{\max(|a|,|b|)}$ целых точек, поэтому

$$|S_0(f)| = |\{v_1 \pm (b, -a), v_2 \pm (b, -a)\} \cap M_0(f)| \\ + |\{(x_1, x_2) \in M_0(f) \mid |ax_1 + bx_2 + c| = 1\}| \leqslant 2 + \frac{2n}{\max(|a|, |b|)}.$$

Отсюда $|S(f)| = |S_1(f)| + |S_0(f)| \leqslant 4 + \frac{2n}{\max(|a|,|b|)}$. Лемма 4 доказана.

Из (1), (2) и леммы 4 следует

Теорема 2. *Пусть* $f \in \mathfrak{T}(2,n,*)$. *Тогда*

$$\sigma(f, \mathfrak{T}(2,n,*)) = \begin{cases} \Theta(n^2), & |M_1(f)| \leqslant 1, \\ O(n), & |M_1(f)| > 1. \end{cases} \tag{11}$$

4. Длина обучения в классе $\mathfrak{T}(2,n,2)$

Влияние геометрических свойств функции на её длину обучения представляет самостоятельный интерес. В [9] рассматриваются k-пороговые функции, для которых $P(f)$ имеет ненулевую площадь. Приводится зависимость длины обучения функции f относительно класса $\mathfrak{T}(2,n,*)$ от таких свойств $P(f)$, как периметр и наименьший угол при вершине (см. лемму 1).

В данной статье получим аналогичные оценки для 2-пороговой функции относительно класса $\mathfrak{T}(2,n,2)$, используя характеристики функции. Напомним, что лемма 3 ограничивает длину обучения большого подмножества 2-пороговых функций константой 9, поэтому такие функции можно исключить из рассмотрения. В лемме 5 доказывается, что все функции, не удовлетворяющие лемме 3, за исключением некоторых пограничных случаев, имеют общее свойство, а именно, могут быть заданы системой пороговых неравенств такой, что соответствующие им пороговые прямые пересекаются внутри $\mathrm{Conv}(E_n^2)$.

Лемма 5. *Пусть* $f \in \mathfrak{T}(2,n,2)$, $B(E_n^2) \cap M_1(f) \neq \varnothing$ *и существует не менее двух различных множеств пороговых функций, определяющих* f. *Тогда найдутся пороговые неравенства, определяющие* f, *такие, что соответствующие пороговые прямые пересекаются в* $\mathrm{Conv}(E_n^2)$.

Доказательство. Заметим, что для того, чтобы показать существование пороговых неравенств, пороговые прямые которых пересекаются в $\mathrm{Conv}(E_n^2)$, достаточно доказать, что существует определяющее множество пороговых функций, имеющих общий нуль. Пусть это не так и для произвольного определяющего множества $\{f_1, f_2\}$ выполнено равенство

$$M_0(f_1) \cap M_0(f_2) = \varnothing. \tag{12}$$

Так как $B(E_n^2) \cap M_1(f) \neq \varnothing$, имеем $S(f_1) \cap M_0(f_2) \neq \varnothing$ или $S(f_2) \cap M_0(f_1) \neq \varnothing$ по лемме 2. Не уменьшая общности, будем считать $S(f_1) \cap$

$M_0(f_2) \neq \varnothing$. Тогда из (12) следует, что $S_1(f_1) \cap M_0(f_2) \neq \varnothing$. Для произвольной точки $x \in S_1(f_1) \cap M_0(f_2)$ рассмотрим функцию $f_1' \in \mathfrak{T}(2, n)$ такую, что $f_1(y) = f_1'(fy)$ и $f_1(x) \neq f_1'(x)$ для всех $y \in E_n^2 \setminus \{x\}$. Множество $\{f_1', f_2\}$ является определяющим для f. Кроме того, $S_0(f_1') \cap M_0(f_2) \neq \varnothing$. Лемма 5 доказана.

Существование пересекающихся внутри $\mathrm{Conv}\big(E_n^2\big)$ пороговых прямых, соответствующих 2-пороговой функции f, можно использовать для оценки длины обучения f в зависимости от угла между пороговыми прямыми.

Введём некоторые вспомогательные обозначения и формальное определение угла между пороговыми прямыми. Для каждой пары нуля и единицы некоторой k-пороговой функции f очевидно, что через отрезок с концами в этих точках проходит как минимум одна из пороговых прямых для каждого определяющего множества пороговых неравенств. Рассмотрим пары нуля и единицы функции f, лежащих на границе E_n^2 на расстоянии 1 друг от друга. Перенумеруем их произвольным образом, и для каждой i-й пары обозначим нуль функции через $b_i(f)$, а единицу — через $v_{b_i}(f)$. Таким образом, $b_i(f)v_{b_i}(f)$ — отрезок длины 1 с концами в целочисленных точках, лежащий на границе $\mathrm{Conv}\big(E_n^2\big)$ и пересекающий как минимум одну из пороговых прямых, определяющих функцию f, в какой-либо точке, отличной от $b_i(f)$.

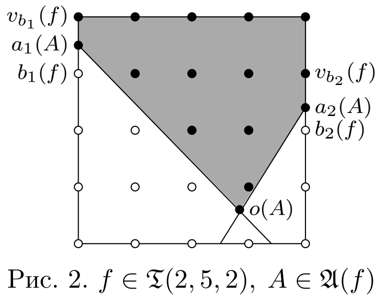

Рис. 2. $f \in \mathfrak{T}(2, 5, 2)$, $A \in \mathfrak{A}(f)$

Пусть $\mathfrak{A}(f)$ — множество всех систем из двух пороговых неравенств, определяющих функцию f. Множество решений некоторой системы неравенств $A \in \mathfrak{A}(f)$, пересеченное с квадратом $\mathrm{Conv}\big(E_n^2\big)$, образует выпуклый многоугольник (серая область на рис. 2). Через $a_i(A)$ обозначим ближайшую к $v_{b_i}(f)$ точку, в которой пересекаются отрезок $b_i(f)v_{b_i}(f)$ и одна из пороговых прямых, соответствующих системе A. Через $o(A)$ обозначим точку пересечения пороговых прямых, если $A \in \mathfrak{A}(f)$ и пороговые прямые пересекаются. На рис. 2 проиллюстрированы введённые понятия на примере функции из класса $\mathfrak{T}(2, 5, 2)$. Угол $\angle a_1(A)o(A)a_2(A)$

назовём *углом между пороговыми прямыми*, соответствующими неравенствам A.

Для k-пороговых функций существует континуальное множество определяющих систем пороговых неравенств, которым будут соответствовать разные углы между пороговыми прямыми, поэтому имеет смысл выбрать некоторое каноническое представление. Например, можно попробовать выбрать такую систему пороговых неравенств, чтобы угол при соответствующих пороговых прямых принимал своё максимальное значение. Максимум по всем углам между пороговыми прямыми, соответствующим неравенствам из $\mathfrak{A}(f)$, в общем случае не достигается. Поэтому будем рассматривать супремум и функцию f', близкую к f, на которой он достигается.

Обозначим $p_{\max}(f) = \sup\limits_{A \in \mathfrak{A}(f)} \angle a_1(A)o(A)a_2(A)$.

Пусть $f \in \mathfrak{T}(2,n,2) \setminus \mathfrak{T}(2,n)$ и $B\big(E_n^2\big) \cap M_1(f) \neq \varnothing$. Если $p_{\max}(f) < \pi$, то *опорной* функцией для f назовём такую функцию $f' \in \mathfrak{T}(2,n,2)$, для которой найдётся множество пороговых неравенств $A' \in \mathfrak{A}(f')$, называемых *опорными* для f и удовлетворяющих условиям:

(i) $f(x) = f'(x)$ для всех $x \in E_n^2 \setminus (o(A')a_1(A') \cup o(A')a_2(A'))$;

(ii) для $i = 1, 2$ существует $x_i \in o(A')a_i(A')$ такой, что

$$o(A')x_i \cap E_n^2 \subset M_1(f), \quad x_i a_i(A') \cap E_n^2 \subset M_0(f),$$

причём $o(A')x_i \cap E_n^2 \neq \varnothing$, $x_i a_i(A') \cap E_n^2 \neq \varnothing$.

(iii) $\angle a_1(A')o(A')a_2(A') = p_{\max}(f)$.

Возьмём некоторые пороговые прямые функции f и начнём их поворачивать по направлению увеличения угла между ними до тех пор, пока обе пороговые прямые не пересекут единицу и существенный нуль функции. Система пороговых неравенств, соответствующая получившимся пороговым прямым, задаёт опорную функцию f'. Докажем существование опорной функции в случае, когда $p_{\max} < \pi$.

Лемма 6. *Пусть $f \in \mathfrak{T}(2,n,2) \setminus \mathfrak{T}(2,n)$, $M_1(f) \cap B\big(E_n^2\big) \neq \varnothing$ и существует не менее двух различных множеств пороговых функций, определяющих f. Тогда если $p_{\max}(f) < \pi$, то для f существует опорная функция f'.*

Доказательство. Сначала приведём алгоритм построения опорной функции f' (рис. 3), если $p_{\max}(f) < \pi$, а затем докажем, что полученная функция является опорной и на ней достигается $p_{\max}(f)$. Для системы пороговых неравенств $A^i \in \mathfrak{A}(f)$ будем обозначать $a_1^i = a_1(A^i)$, $o^i = o(A^i)$, $a_2^i = a_2(A^i)$.

Шаг 1. Возьмём произвольную пару пороговых неравенств $A^0 \in \mathfrak{A}(f)$, пороговые прямые которых пересекаются внутри $\mathrm{Conv}\big(E_n^2\big)$, и для каждого $j \in \{1, 2\}$ выполним следующие действия.

Шаг 2. Положим $i = 0$.

Шаг 3. Положим $x^i = \arg \min\limits_{x \in \{a_j^i\} \cup M_1(f) \cap m^i} l(o^i x)$, где m^i — прямая, проходящая через отрезок $o^i a_j^i$. Повернём m^i вокруг точки x^i по направлению увеличения угла между пороговыми прямыми на минимальный угол, при котором произойдёт одно из событий шага 4.

Шаг 4. (a) Прямая пересекла точку из $M_1(f) \setminus o^i a_j^i$. Обозначим полученную прямую через m^{i+1}, соответствующую пару пороговых неравенств через A^{i+1}, увеличим i на 1 и вернёмся к шагу 3.

(b) Прямая пересекла точку из $M_0(f)$, в которой неравенство, соответствующее прямой m^i, не выполняется. Неравенство, соответствующее полученной прямой, обозначим через e_j, а соответствующую ему прямую — через m_j.

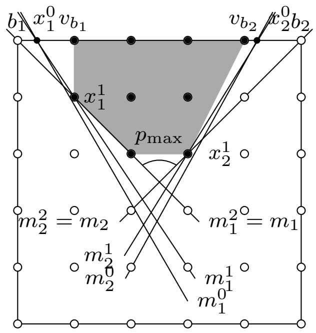

Рис. 3. Пример работы алгоритма,
серая область — $P(f)$

На выходе алгоритма получили множество неравенств $A = \{e_1, e_2\}$, задающее некоторую 2-пороговую функцию f'.

Из построения m_1 и m_2 следует, что $p_{\max}(f) = \pi$ при $m_1 = m_2$.

Пусть $m_1 \neq m_2$. Докажем, что f' является опорной для функции f. Пп. (i) и (ii) определения опорной функции следуют из построения A. Для доказательства п. (iii) достаточно установить два факта. Во-первых, в ходе алгоритма угол между пороговыми прямыми не уменьшается, т. е. для любой входной системы $A^0 \in \mathfrak{A}(f)$ на выходе получаем систему $A \in \mathfrak{A}(f')$ такую, что $\angle a_1(A^0)o(A^0)a_2(A^0) \leqslant \angle a_1(A)o(A)a_2(A)$. Во-вторых, результат работы алгоритма не зависит от входной системы A^0.

Первый факт непосредственно следует из описания алгоритма. Для доказательства второго факта предположим, что для каких-нибудь различных входных систем из $\mathfrak{A}(f)$ алгоритм возвращает различные системы пороговых неравенств A и A'. Обозначим через $\{m_1, m_2\}$ и $\{m_1', m_2'\}$ множества пороговых прямых, соответствующие этим системам неравенств. В каждом из указанных множеств пороговых прямых хотя бы одна прямая должна пересекать отрезок vb_1, пусть это будут m_1 и m_1'. Положим

$$x = \arg \min_{v \in \mathrm{Vert}(P(f)) \cap m_1} l(va_1(A)),\ x' = \arg \min_{v \in \mathrm{Vert}(P(f)) \cap m_1'} l(va_1(A')).$$

Последнее равенство следует из того факта, что неравенства, соответствующие прямым m_1 и m_1', выполняются во всех точках $\mathrm{Vert}(P(f))$. В соответствии с шагом 4(a) алгоритма оба отрезка $a_1(A)x$ и $a_1(A')x$ содержат хотя бы одну точку из $M_0(f)$. Предположим, что $l(a_1(A)b_1) < l(a_1(A')b_1)$. Тогда отрезок $a_1(A')x'$ полностью лежит внутри треугольника $\triangle a_1(A)v_{b_1}x$, чего может не быть, так как треугольник $\triangle a_1(A)v_{b_1}x$ не содержит точек из $M_0(f)$. Аналогично можно показать, что $l(a_1(A)b_1) \not> l(a_1(A')b_1)$. Следовательно, $l(a_1(A)b_1) = l(a_1(A')b_1)$, т. е. $a_1(A) = a_1(A')$, а значит, $m_1 = m_1'$. Лемма 6 доказана.

Лемма 7 связывает минимальный угол при вершине $P(f)$ и $p_{\max}(f)$. В дальнейшем эта связь позволит нам исключить минимальный угол при вершине $P(f)$ из оценки длины обучения 2-пороговой функции.

Лемма 7. *Пусть* $f \in \mathfrak{T}(2, n, 2) \setminus \mathfrak{T}(2, n)$, $M_1(f) \cap B\big(E_n^2\big) \neq \varnothing$, $p_{\max}(f) < \pi$, $\mathcal{S}(P(f)) > 0$ *и для* f *определяющее множество пороговых функций не единственно. Тогда*

$$q_{\min}(P(f)) \geqslant \max\left(p_{\max} - \arcsin\frac{1}{l_1} - \arcsin\frac{1}{l_2}, \frac{1}{4(l_1 + l_2)^2}\right),$$

где $l_i = l(oa_i)$, $o = o(A)$, $a_1 = a_1(A)$, $a_2 = a_2(A)$, A — *пара опорных неравенств для* f *и* $p_{\max} = p_{\max}(f)$.

Доказательство. Разобьём доказательство на две части. Сначала докажем, что

$$q_{\min}(P(f)) \geqslant \frac{1}{4(l_1 + l_2)^2}.$$

Заметим, что $q(v, P(f)) \geqslant \angle v_{b_1} v v_{b_2}$ для любой вершины $v \in \mathrm{Vert}(P(f)) \setminus B\big(E_n^2\big)$. В зависимости от мощности $\mathrm{Vert}\big(\mathrm{Conv}\big(E_n^2\big)\big) \cap P(f)$ можно выделить четыре случая.

Случай 1: $\mathrm{Vert}\big(\mathrm{Conv}\big(E_n^2\big)\big) \cap P(f) = \{x, y\}$. Тогда для любой вершины $v \in \mathrm{Vert}(P(f)) \setminus B\big(E_n^2\big)$ имеем

$$\angle v_{b_1} v v_{b_2} \geqslant \angle xvy \geqslant \min_{z \in E_n^2} (\angle xzy) \geqslant \frac{\pi}{4}.$$

Случай 2: $\mathrm{Vert}\big(\mathrm{Conv}\big(E_n^2\big)\big) \setminus P(f) = \{x\}$. Тогда для любой вершины $v \in \mathrm{Vert}(P(f)) \setminus B\big(E_n^2\big)$ по утверждению 3 получаем

$$\angle v_{b_1} v v_{b_2} \geqslant \angle v_{b_1} x v_{b_2} = \frac{\pi}{2}.$$

Случай 3: $\mathrm{Vert}\big(\mathrm{Conv}\big(E_n^2\big)\big) \cap P(f) = \{x\}$. Для любых трёх вершин $v_1, v_2, v_3 \in \mathrm{Vert}(P(f))$ треугольник $\triangle v_1 v_2 v_3$ содержится в четырёхугольнике $P_4 = \mathrm{Conv}(\{a_1, o, a_2, x\})$. Отсюда по утверждению 1 выводим

$$\min_{i=1,2,3} q(v_i, P(f)) \geqslant \arcsin \frac{1}{\mathrm{Diam}(\triangle v_1 v_2 v_3)^2} \geqslant \arcsin \frac{1}{\mathrm{Diam}(P_4)^2}.$$

Далее,

$$\mathrm{Diam}(P_4) \leqslant \mathrm{Diam}(\triangle a_1 o a_2) + \mathrm{Diam}(\triangle a_1 x a_2) \leqslant (l_1 + l_2) + l(a_1 a_2) \leqslant 2(l_1 + l_2).$$

Таким образом,

$$q_{\min}(P(f)) \geqslant \min_{v_1, v_2, v_3 \in \mathrm{Vert}(P(f))} \angle v_1 v_2 v_3 \geqslant \arcsin \frac{1}{4(l_1 + l_2)^2} \geqslant \frac{1}{4(l_1 + l_2)^2}.$$

Случай 4: $\mathrm{Vert}\big(\mathrm{Conv}\big(E_n^2\big)\big) \cap P(f) = \varnothing$. В этом случае рассуждениями, аналогичными случаю 3, легко показать, что $q_{\min}(P(f)) \geqslant \frac{1}{(l_1 + l_2)^2}$.

Резюмируя все четыре случая, получаем

$$q_{\min}(P(f)) \geqslant \frac{1}{4(l_1 + l_2)^2}. \tag{13}$$

Обозначим $v_{a_i} = \arg \min\limits_{v \in P(f) \cap oa_i} l(ov)$. По лемме 6 такая точка найдётся, причём $v_{a_i} \notin B\big(E_n^2\big)$. Не уменьшая общности, будем считать, что при обходе вершин $P(f)$ по часовой стрелке обозначенные вершины выстроены в порядке $v_{a_1}, v_{b_1}, v_{b_2}, v_{a_2}$ (рис. 4).

Обозначим через h_1 проекцию точки v_{b_1} на прямую, проходящую через отрезок oa_1. Треугольник $\triangle v_{a_1} v_{b_1} h_1$ прямоугольный с гипотенузой $v_{a_1} v_{b_1}$, причём все вершины $P(f)$, встречающиеся на пути от v_{a_1} до v_{b_1}

при обходе по часовой стрелке, находятся в этом треугольнике. При каждой такой вершине v в $P(f)$ по утверждению 3 имеем

$$q(v, P(f)) \geqslant \angle v_{a_1} v v_{b_1} \geqslant \angle v_{a_1} h_1 v_{b_1} \geqslant \frac{\pi}{2}. \tag{14}$$

Аналогично показывается, что неравенство (14) справедливо для всех вершин, встречающихся на пути от v_{a_2} до v_{b_2} при обходе против часовой стрелки.

Все вершины $P(f)$, находящиеся по пути обхода по часовой стрелке от v_{a_2} до v_{a_1}, если таковые есть, находятся в треугольнике $\triangle v_{a_1} v_{a_2} o$. По утверждению 3

$$\min_{v \in \operatorname{Vert} P(f) \cap \triangle v_{a_1} v_{a_2} o \setminus \{v_{a_1}, v_{a_2}\}} q(v, P(f)) \geqslant \angle v_{a_1} o v_{a_2} = p_{\max}.$$

Мы оценили снизу углы при всех вершинах в $P(f) \setminus B\big(E_n^2\big)$ за исключением вершин v_{a_1} и v_{a_2}, поэтому рассмотрим эти вершины и углы при них. Пусть v_{a_1} и v_{a_2} совпадают, т. е. $v_{a_1} = v_{a_2} = o$. Тогда

$$q(o, P(f)) \geqslant \angle v_{b_1} o v_{b_2} = p_{\max} - \angle a_1 o v_{b_1} - \angle a_2 o v_{b_2}.$$

Из того, что $l(a_i v_{b_i}) \leqslant 1$, следует, что $\angle a_i o v_{b_i} \leqslant \arcsin \frac{1}{l_i}$, $i = 1, 2$, поэтому

$$q(o, P(f)) \geqslant p_{\max} - \arcsin \frac{1}{l_1} - \arcsin \frac{1}{l_2}. \tag{15}$$

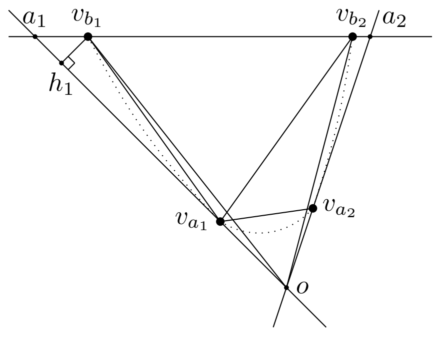

Рис. 4. Пунктирной линией схематично обозначена граница $P(f)$

Пусть $v_{a_1} \neq v_{a_2}$ (см. рис. 4). Рассмотрим угол при вершине v_{a_1}:

$$q(v_{a_1}, P(f)) \geqslant \angle v_{b_1} v_{a_1} v_{a_2} > \angle v_{b_1} v_{a_1} v_{b_2}. \tag{16}$$

Пусть $\angle v_{b_1} v_{a_1} v_{b_2} \leqslant \frac{\pi}{6}$. Угол $\angle v_{b_1} v_{a_1} a_1$ является наименьшим углом в треугольнике $\triangle v_{b_1} v_{a_1} a_1$, поэтому $\angle v_{b_1} v_{a_1} a_1 \leqslant \frac{\pi}{3}$ и $\angle v_{b_2} v_{a_1} a_1 \leqslant \frac{\pi}{2}$. Тогда по

теореме синусов, применённой к треугольникам $\triangle a_1 v_{a_1} v_{b_1}$ и $\triangle a_1 o v_{b_2}$, получаем

$$l(v_{b_2} v_{a_1}) \leqslant l(v_{b_2} o).$$

Отсюда по теореме синусов в треугольниках $\triangle v_{b_1} o v_{b_2}$ и $\triangle v_{b_2} v_{a_1} v_{b_1}$ получаем $\sin \angle v_{b_2} v_{a_1} v_{b_1} > \sin \angle v_{b_1} o v_{b_2}$. Так как $\angle v_{b_2} v_{a_1} v_{b_1} < \frac{\pi}{6}$, справедливо неравенство $\angle v_{b_1} v_{a_1} v_{b_2} > \angle v_{b_1} o v_{b_2}$. Из (15) и (16) следует, что

$$q_{\min}(P(f)) \geqslant p_{\max} - \arcsin \frac{1}{l_1} - \arcsin \frac{1}{l_2}.$$

Учитывая (13), получаем утверждение леммы. Лемма 7 доказана.

Теорема 3 даёт оценку мощности тупикового разрешающего множества функции $f \in \mathfrak{T}(2, n, 2)$ в зависимости от $p_{\max}(f)$ с помощью двух дополнительных параметров l_1, l_2, введённых в лемме 7.

Теорема 3. *Пусть $f \in \mathfrak{T}(2, n, 2) \setminus \mathfrak{T}(2, n)$, $B\big(E_n^2\big) \cap M_1(f) \neq \varnothing$ и $M_1(f) \nsubseteq B\big(E_n^2\big)$. Пусть также для f существует как минимум два различных определяющих множества пороговых функций и $p_{\max} < \pi$, где $p_{\max} = p_{\max}(f)$. Тогда*

$$\sigma(f, \mathfrak{T}(2, n, 2)) < 18 + \frac{110}{\pi - p_{\max}} + 4\pi \max \left(\frac{1}{\max\left(p_{\max} - \arcsin \frac{1}{l_1} - \arcsin \frac{1}{l_2}, \frac{1}{4(l_1 + l_2)^2}\right)}, 1 + \frac{4\sqrt{2} + 2}{\pi - p_{\max}} \right),$$

где $l_i = l(oa_i)$, $o = o(A)$, $a_1 = a_1(A)$, $a_2 = a_2(A)$, A — пара опорных неравенств для f, $p_{\max} = p_{\max}(f)$.

Доказательство. Множество $\Delta P(f) \cup \mathrm{Vert}(P(f))$ по теореме 1 и утверждению 6 разрешающее для f относительно класса $\mathfrak{T}(2, n, *)$, а следовательно, и относительно класса $\mathfrak{T}(2, n, 2)$. В ходе доказательства выделим некоторое множество $D \subset E_n^2$ так, что функция f', заданная на нём и совпадающая всюду с f, удовлетворяет условиям следствия 2. В подобранном таким образом D существует разрешающее множество f' мощности не больше 9, по которому также можно однозначно восстановить функцию f на множестве $(\Delta P(f) \cup \mathrm{Vert}(P(f))) \cap D$. Объединение этих 9 или менее точек и $(\Delta P(f) \cup \mathrm{Vert}(P(f))) \setminus D$ будет разрешающим множеством f относительно $\mathfrak{T}(2, n, 2)$, и останется лишь оценить его мощность.

Проведём касательную к $P(f)$, наиболее близкую к точке o и перпендикулярную биссектрисе угла $\angle a_1 o a_2$, и обозначим её через c (рис. 5).

Вершины $P(f)$, лежащие на c, обозначим через v_{c_1} и v_{c_2} и заметим, что если касание с $P(f)$ произошло в единственной точке, то $v_{c_1} \equiv v_{c_2}$. Касательная c разделяет $\mathrm{Conv}\bigl(E_n^2\bigr)$ на два выпуклых множества, одно из которых полностью включает $P(f)$. Возьмём в качестве D пересечение этого множества с E_n^2 и заметим, что $(\Delta P(f) \cup \mathrm{Vert}(P(f))) \setminus D = \Delta P(f) \setminus D$.

Функция f', заданная на D и такая, что $f(x) = f'(x)$ для всех $x \in D$, удовлетворяет условиям следствия 2 и имеет разрешающее множество мощности не больше 9. Объединение этого множества и $\Delta P(f) \setminus D$ будет разрешающим множеством функции f, а значит,

$$\sigma(f, \mathfrak{T}(2, n, 2)) \leqslant 9 + \bigl|(\Delta P(f) \setminus D) \cap E_n^2\bigr|.$$

Так как все целые точки в $\Delta P(f)$ находятся на границе $P'(f)$, число целых точек в $\Delta P(f) \setminus D$ может быть оценено следующим образом:

$$\bigl|(\Delta P(f) \setminus D) \cap E_n^2\bigr| \leqslant l(b(P'(f)) \setminus \mathrm{Conv}(D)) + 1.$$

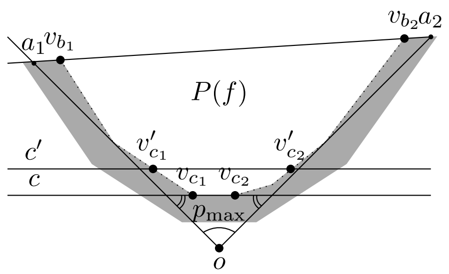

Рис. 5. Область $\Delta P(f)$ закрашена серым цветом

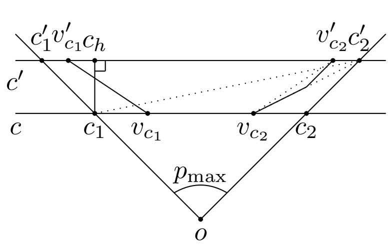

Рис. 6. Касательная c и параллельная ей прямая c'

Через $P(f)$ проведём прямую c', параллельную касательной c и удалённую от неё на расстояние 1 (рис. 6). Точки пересечения прямой c' с $b(P(f))$ обозначим через v'_{c_1} и v'_{c_2}. Через C обозначим множество точек из $P(f)$, лежащих между прямыми c и c' и на них. Через C' обозначим расширение C. Нетрудно видеть, что $b(P'(f)) \setminus \mathrm{Conv}(D) \subset b(C')$, поэтому

$$l(b(P'(f)) \setminus \mathrm{Conv}(D)) \leqslant \mathcal{P}(C') \leqslant 5\mathcal{P}(C) + \frac{4}{\sin \frac{q_{\min}(C)}{2}} + 8,$$

причём

$$q_{\min}(C) \geqslant \min(q_{\min}(P(f)), q(v'_{c_1}, C), q(v'_{c_2}, C)). \tag{17}$$

Таким образом,

$$\sigma(f, \mathfrak{T}(2, n, 2)) \leqslant 18 + 5\mathcal{P}(C) + \frac{4}{\sin \frac{q_{\min}(C)}{2}}. \tag{18}$$

Для начала оценим $\mathcal{P}(C)$, C — выпуклый многоугольник, вписанный в равнобедренную трапецию, заключенную между c и c', на которых лежат основания трапеции, и прямыми, проходящими через отрезки oa_1 и oa_2. Обозначим через c_1, c_2, c'_1, c'_2 точки пересечения c и oa_1, c и oa_2, c' и oa_1, c' и oa_2 соответственно. Так как C вписан в трапецию $c_1c_2c'_2c'_1$, $\mathcal{P}(C) \leqslant \mathcal{P}(c_1c_2c'_2c'_1)$. Трапеция $c_1c_2c'_2c'_1$ имеет высоту 1 и углы при большем основании, равные $\frac{\pi-p_{\max}}{2}$, поэтому

$$\mathcal{P}(C) \leqslant \mathcal{P}(c_1c_2c'_2c'_1) = 2l(c_1c_2) + \frac{2}{\sin\frac{\pi-p_{\max}}{2}} + 2\operatorname{ctg}\frac{\pi-p_{\max}}{2}.$$

Для оценки $l(c_1c_2)$ сверху заметим, что в треугольнике $\triangle c_1oc_2$ нет внутренних целых точек и $\angle oc_1c_2 = \angle oc_2c_1 = \frac{\pi-p_{\max}}{2}$. По утверждению 2 получаем

$$l(c_1c_2) \leqslant \sqrt{2}\operatorname{ctg}\frac{\pi-p_{\max}}{4}. \tag{19}$$

Таким образом,

$$\mathcal{P}(C) \leqslant 2\sqrt{2}\operatorname{ctg}\frac{\pi-p_{\max}}{4} + \frac{2}{\sin\frac{\pi-p_{\max}}{2}} + 2\operatorname{ctg}\frac{\pi-p_{\max}}{2}.$$

Поскольку $0 < p_{\max} < \pi$, имеем $0 < \frac{\pi-p_{\max}}{4} < \frac{\pi-p_{\max}}{2} < \frac{\pi}{2}$. Так как для любого $0 < x < \frac{\pi}{2}$ выполнено (4) и (5), то

$$\mathcal{P}(C) \leqslant \frac{8\sqrt{2}}{\pi-p_{\max}} + \frac{2\pi}{\pi-p_{\max}} + \frac{4}{\pi-p_{\max}} < \frac{22}{\pi-p_{\max}}. \tag{20}$$

Теперь оценим сверху третье слагаемое в сумме (18), воспользовавшись неравенством (17):

$$\frac{1}{\sin\frac{q_{\min}(C)}{2}} \leqslant \max\left(\frac{1}{\sin\frac{q_{\min}(P(f))}{2}}, \frac{1}{\sin\frac{q(v'_{c_1},C)}{2}}, \frac{1}{\sin\frac{q(v'_{c_2},C)}{2}}\right). \tag{21}$$

Угол $q_{\min}(P(f))$ оценен снизу в лемме 7, поэтому по (5) имеем

$$\frac{1}{\sin\frac{q_{\min}(P(f))}{2}} < \frac{\pi}{\max\left(p_{\max} - \arcsin\frac{1}{l_1} - \arcsin\frac{1}{l_2}, \frac{1}{4(l_1+l_2)^2}\right)}. \tag{22}$$

Далее оценим снизу углы $q(v'_{c_1}, C)$ и $q(v'_{c_2}, C)$. Из утверждения 3 выводим

$$q(v'_{c_2}, C) \geqslant \angle c'_1v'_{c_2}v_{c_2} \geqslant \angle c'_1c'_2v_{c_2} \geqslant \angle c'_1c'_2c_1.$$

Пусть c_h — проекция точки c_1 на прямую c'. Угол $\angle c_1'c_2'c_1$ можно вычислить по катетам треугольника $\triangle c_1c_hc_2'$:

$$\angle c_1'c_2'c_1 = \operatorname{arctg}\frac{l(c_1c_h)}{l(c_hc_2')} = \operatorname{arctg}\frac{1}{\left(l(c_1c_2) + \operatorname{ctg}\frac{\pi-p_{\max}}{2}\right)}. \tag{23}$$

Последнее равенство следует из того, что $l(c_1c_h) = 1$, $l(c_1'c_2') = l(c_1c_2) + 2l(c_1'c_h)$ и $l(c_1'c_h) = \operatorname{ctg}\frac{\pi-p_{\max}}{2}$. Далее, используя неравенство (5), получим

$$\frac{1}{\sin\frac{q(v_{c_2}',C)}{2}} \leqslant \frac{1}{\sin\frac{\angle c_1'c_2'c_1}{2}} < \frac{\pi}{\angle c_1'c_2'c_1} < \frac{\pi}{\sin\angle c_1'c_2'c_1}.$$

Так как $\angle c_1'c_2'c_1 < \frac{\pi}{2}$, то

$$\frac{\pi}{\sin\angle c_1'c_2'c_1} < \frac{\pi(\sin\angle c_1'c_2'c_1 + \cos\angle c_1'c_2'c_1)}{\sin\angle c_1'c_2'c_1} = \pi + \pi\operatorname{ctg}\angle c_1'c_2'c_1.$$

Подставляя $\operatorname{ctg}\angle c_1'c_2'c_1$ из неравенства (23), приходим к равенству

$$\pi + \pi\operatorname{ctg}\angle c_1'c_2'c_1 = \pi\left(1 + l(c_1c_2) + \operatorname{ctg}\frac{\pi - p_{\max}}{2}\right).$$

Из неравенства (19) имеем

$$\pi\left(1 + l(c_1c_2) + \operatorname{ctg}\frac{\pi - p_{\max}}{2}\right) \leqslant \pi\left(1 + \sqrt{2}\operatorname{ctg}\frac{\pi - p_{\max}}{4} + \operatorname{ctg}\frac{\pi - p_{\max}}{2}\right)$$

и, применив неравенство (4) к последнему выражению, получим

$$\frac{1}{\sin\frac{q(v_{c_2}',C)}{2}} < \pi\left(1 + \frac{4\sqrt{2}+2}{\pi - p_{\max}}\right).$$

Аналогично оценивается выражение $\dfrac{1}{\sin\frac{q(v_{c_1}',C)}{2}}$.

Перепишем неравенство (21), воспользовавшись полученным результатом и (22):

$$\frac{1}{\sin\frac{q_{\min}(C)}{2}} < \pi\max\left(\frac{1}{\max\left(p_{\max} - \arcsin\frac{1}{l_1} - \arcsin\frac{1}{l_2}, \frac{1}{4(l_1+l_2)^2}\right)},\right.$$
$$\left. 1 + \frac{4\sqrt{2}+2}{\pi - p_{\max}}\right). \tag{24}$$

В силу последнего неравенства и (20) можем переписать равенство (18) в следующем виде:

$$\sigma(f, \mathfrak{T}(2,n,2)) < 18 + \frac{110}{\pi - p_{\max}} \\ + 4\pi \max\left(\frac{1}{\max\left(p_{\max} - \arcsin\frac{1}{l_1} - \arcsin\frac{1}{l_2}, \frac{1}{4(l_1+l_2)^2}\right)}, 1 + \frac{4\sqrt{2}+2}{\pi - p_{\max}}\right).$$

Замечание 2. Параметры l_1, l_2 как длины отрезков, заключённых в $\mathrm{Conv}(E_n^2)$ и содержащих как минимум две целые точки каждый, имеют следующее ограничение:

$$1 \leqslant l_i \leqslant \sqrt{2}n \quad \text{для } i = 1, 2. \tag{25}$$

Так как угол $p_{\max}(f)$ находится между отрезками, содержащими не менее двух целых точек каждый и имеющими ограничение длины (25), используя утверждение 1, нетрудно показать, что

$$\frac{1}{2n^2} \leqslant p_{\max}(f) \leqslant \pi - \frac{1}{n^2}.$$

Применим указанные оценки к правой части неравенства из утверждения теоремы 3, чтобы сравнить её с функцией от n:

$$\sigma(f, \mathfrak{T}(2,n,2)) < 18 + \frac{110}{\pi - p_{\max}} \\ + 4\pi \max\left(\frac{1}{\max\left(p_{\max} - \arcsin\frac{1}{l_1} - \arcsin\frac{1}{l_2}, \frac{1}{4(l_1+l_2)^2}\right)},\right. \\ \left. 1 + \frac{4\sqrt{2}+2}{\pi - p_{\max}}\right) = O(n^2).$$

Полученная оценка не улучшает результат теоремы 2 в общем случае. При этом теорема 3 даёт хорошую верхнюю оценку длины обучения 2-пороговой функции f относительно класса $\mathfrak{T}(2,n,2)$ в случае, когда угол $p_{\max}(f)$ ограничен снизу и сверху некоторой константой. Формально данное утверждение сформулировано в следствии 3.

Следствие 3. *Пусть $f \in \mathfrak{T}(2,n,2)$ удовлетворяет условиям теоремы 3 и $p_{\max} = p_{\max}(f) \in (\varepsilon, \pi - \varepsilon)$ для некоторого $0 < \varepsilon < \frac{\pi}{2}$. Тогда*

$$\sigma(f, \mathfrak{T}(2,n,2)) = \begin{cases} O\left(\frac{1}{\varepsilon}\right), & l_1, l_2 \geqslant \frac{2\pi}{p_{\max}}, \\ O\left(\frac{1}{\varepsilon^2}\right), & l_1, l_2 < \frac{2\pi}{p_{\max}}. \end{cases}$$

Доказательство. Пусть $l_1, l_2 \geqslant \frac{2\pi}{p_{\max}}$. Тогда с использованием неравенства (5) для $i = 1, 2$ получаем

$$\arcsin \frac{1}{l_i} < \arcsin \frac{p_{\max}}{2\pi} < \frac{p_{\max}}{4},$$

откуда

$$p_{\max} - \arcsin \frac{1}{l_1} - \arcsin \frac{1}{l_2} > \frac{p_{\max}}{2}$$

и

$$\max\left(p_{\max} - \arcsin \frac{1}{l_1} - \arcsin \frac{1}{l_2}, \frac{1}{4(l_1 + l_2)^2}\right) > \frac{p_{\max}}{2} > \frac{\varepsilon}{2}.$$

Отсюда

$$\begin{aligned}\sigma(f, \mathfrak{T}(2, n, 2)) &< 18 + \frac{110}{\varepsilon} + 4\pi \max\left(\frac{2}{\varepsilon}, 1 + \frac{4\sqrt{2} + 2}{\varepsilon}\right)\\ &= 18 + 4\pi + \frac{110 + 16\pi\sqrt{2} + 8\pi}{\varepsilon}.\end{aligned}$$

Теперь пусть $l_1, l_2 < \frac{2\pi}{p_{\max}}$. Тогда

$$\frac{1}{4(l_1 + l_2)^2} > \frac{p_{\max}^2}{32\pi^2},$$

а значит,

$$\max\left(p_{\max} - \arcsin \frac{1}{l_1} - \arcsin \frac{1}{l_2}, \frac{1}{4(l_1 + l_2)^2}\right) > \frac{p_{\max}^2}{32\pi^2} > \frac{\varepsilon^2}{32\pi^2}$$

и

$$\begin{aligned}\sigma(f, \mathfrak{T}(2, n, 2)) &< 18 + \frac{110}{\varepsilon} + 4\pi \max\left(\frac{32\pi^2}{\varepsilon^2}, 1 + \frac{4\sqrt{2} + 2}{\varepsilon}\right)\\ &< 18 + \frac{110}{\varepsilon} + 4\pi\left(\frac{32\pi^2}{\varepsilon^2} + 1 + \frac{4\sqrt{2} + 2}{\varepsilon}\right)\\ &= 18 + 4\pi + \frac{110 + 16\pi\sqrt{2} + 8\pi}{\varepsilon} + \frac{32\pi^2}{\varepsilon^2}.\end{aligned}$$

Следствие 3 доказано.

Автор выражает признательность рецензенту за внимательное прочтение статьи и ценные замечания, которые значительно улучшили изложение результатов работы.

ЛИТЕРАТУРА

1. **Золотых Н. Ю., Шевченко В. Н.** Об оценке сложности расшифровки пороговых функций k-значной логики // Журн. вычисл. математики и мат. физики. 1999. Т. 39, №. 2. С. 346–352.
2. **Шевченко В. Н., Золотых Н. Ю.** О сложности расшифровки пороговых функций k-значной логики // Докл. АН. 1998. Т. 362, № 5. С. 606–608.
3. **Alekseyev M. A., Basova M. G., Zolotykh N. Yu.** On the minimal teaching sets of two-dimensional threshold functions // SIAM J. Discrete Math. 2015. Vol. 29, No. 1. P. 157–165.
4. **Anthony M., Brightwell G., Shawe-Taylor J.** On specifying Boolean functions by labelled examples // Discrete Appl. Math. 1995. Vol. 61, No. 1. P. 1–25.
5. **Bultman W. J., Maass W.** Fast identification of geometric objects with membership queries // Inf. Comput. 1995. Vol. 118. P. 48–64.
6. **Chirkov A. Yu., Zolotykh N. Yu.** On the number of irreducible points in polyhedra // Graphs Comb. 2016. Vol. 32, No. 5. P. 1789–1803.
7. **Shevchenko V. N., Zolotykh N. Yu.** Lower bounds for the complexity of learning half-spaces with membership queries // Algorithmic Learning Theory. Proc. 9th Int. Conf. (Otzenhausen, Germany, Oct. 8–10, 1998). Berlin: Springer, 1998. P. 61–71. (Lect. Notes Comput. Sci.; Vol. 1501).
8. **Trainin J.** An elementary proof of Pick's theorem // Math. Gaz. 2007. Vol. 91. P. 536–540.
9. **Zamaraeva E.** On teaching sets of k-threshold functions // ArXiv e-prints (arXiv:1502.04340). 2015.

Замараева Елена Михайловна

Статья поступила
31 августа 2015 г.

Исправленный вариант —
2 августа 2016 г.

DISKRETNYI ANALIZ I ISSLEDOVANIE OPERATSII
/DISCRETE ANALYSIS AND OPERATIONS RESEARCH/
January–March 2017. Volume 24, No. 1. P. 31–55

UDC 519.715 DOI: 10.17377/daio.2017.24.508

ON TEACHING SETS FOR 2-THRESHOLD FUNCTIONS OF TWO VARIABLES

E. M. Zamaraeva
Lobachevsky State University,
23 Gagarin Ave., 603950 Nizhny Novgorod, Russia
E-mail: elena.zamaraeva@gmail.com

Abstract. We consider k-threshold functions of n variables, i. e. the functions representable as the conjunction of k threshold functions. For $n = 2$, $k = 2$, we give upper bounds for the cardinality of the minimal teaching set depending on the various properties of the function. Illustr. 6, bibliogr. 9.

Keywords: machine learning, threshold function, teaching dimension, teaching set.

REFERENCES

1. **N. Yu. Zolotykh** and **V. N. Shevchenko,** Estimating the complexity of deciphering a threshold function in a k-valued logic, *Zh. Vychisl. Mat. Mat. Fiz.,* **39**, No. 2, 346–352, 1999 [Russian]. Translated in *Comput. Math. Math. Phys.,* **39**, No. 2, 328–334, 1999.
2. **V. N. Shevchenko** and **N. Yu. Zolotykh,** On the complexity of deciphering the threshold functions of k-valued logic, *Dokl. Akad. Nauk,* **362**, No. 5, 606–608, 1998 [Russian]. Translated in *Dokl. Math.,* **58**, No. 2, 268–270, 1998.
3. **M. A. Alekseyev, M. G. Basova,** and **N. Yu. Zolotykh,** On the minimal teaching sets of two-dimensional threshold functions, *SIAM J. Discrete Math.,* **29**, No. 1, 157–165, 2015.
4. **M. Anthony, G. Brightwell,** and **J. Shawe-Taylor,** On specifying Boolean functions by labelled examples, *Discrete Appl. Math.,* **61**, No. 1, 1–25, 1995.
5. **W. J. Bultman** and **W. Maass,** Fast identification of geometric objects with membership queries, *Inf. Comput.,* **118**, No. 1, 48–64, 1995.
6. **A. Yu. Chirkov** and **N. Yu. Zolotykh,** On the number of irreducible points in polyhedra, *Graphs Comb.,* **32**, No. 5, 1789–1803, 2016.
7. **V. N. Shevchenko** and **N. Yu. Zolotykh,** Lower bounds for the complexity of learning half-spaces with membership queries, in *Algorithmic Learning Theory* (Proc. 9th Int. Conf., Otzenhausen, Germany, Oct. 8–10, 1998), pp. 61–71, Springer, Berlin, 1998 (Lect. Notes Comput. Sci., Vol. 1501).

8. **J. Trainin,** An elementary proof of Pick's theorem, *Math. Gaz.,* **91**, No. 522, 536–540, 2007.
9. **E. M. Zamaraeva** On teaching sets of k-threshold functions, 2015 (Cornell Univ. Libr. e-Print Archive, arXiv:1502.04340).

Elena M. Zamaraeva

Received
31 August 2015
Revised
2 August 2016