Вопросы полноты в классе кусочно-линейных непрерывных функций.

А. Н. Кан

В статье рассматривается класс всех двуместных кусочно-линейных непрерывных функций. Доказывается что данный класс лежит в классе согласованных функций. Найден критерый полноты в этом классе.

Ключевые слова: Класс кусочно-линейных функций, класс кусочно-линейных непрерывных функций, класс согласованных функций, класс финитно-параллельных непрерывных функций, функция Хэвисайда, операции суперпозиции, вектор сигнатуры.

Введение.

В настоящей работе рассматривается класс CPL кусочно-линейных непрерывных функций. Данный класс является подклассом PL кусочно-линейных функций. Кусочно-линейные функции были изучены в работе [1]. Был выделен класс PP кусочно-параллельных функций, в котором был найден предполный класс C-финитно-линейных функций ΦL. Был сформулирован критерий полноты в классе кусочно-параллельных функций. В работе [4] был получен критерий полноты в классе кусочно-линейных функций. Были найдены три замкнутых класса: класс финитных функций Φ, класс кусочно-линейных непрерывных функций CPL и класс согласованных функций P. Данные классы образуют критериальную систему в классе кусочно-линейных функций. Каждый из приведенных классов может обладать "хорошими"свойствами, имеющие значение в реализации нейронной сети. Поэтому отдельно изучается класс кусочно-линейных непрерывных функций. Данный класс был рассмотрен в работе [3]. Оказалось что все кусочно-линейные непрерывные функции зависящие от одной переменной $CPL^{(1)}$ можно получить из функции модуля и линейных функций по операциям суперпозиции. Хотелось бы обобщить результат прошлой работы на все кусочно-линейные непрерывные функции. В настоящей работе была поставлена задача полноты в классе кусочно-линейных непрерывных функций зависящих от двух переменных.

Основные понятия и определения.

Определение 1. *Функция $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ называется линейной, если найдутся $\overline{a} \in \mathbb{R}^n$ и $c \in \mathbb{R}$, такие что $f(\overline{x}) = \overline{x} * \overline{a} + c$, где под операцией "$*$" понимается скалярное произведение векторов. Множество всех линейных функций обозначим через L.*

Пусть l_i - гиперплоскость, задаваемая уравнением $\overline{x} * \overline{a_i} + c_i = 0, a_i \in \mathbb{R}^n \backslash \{\overline{0}\}, c_i \in \mathbb{R}, i = 1, .., k$. Для каждой точки $\overline{x} \in \mathbb{R}^n$ рассмотрим вектор $\sigma(\overline{x}) = (\sigma_1, .., \sigma_k)$ с компонентами из множества $\{-1, 0, 1\}, \sigma_i = sgn(\overline{x} * \overline{a_i} + c_i)$, где

$$sgn(b) = \begin{cases} -1, & \text{если } b < 0 \\ 0, & \text{если } b = 0 \\ 1, & \text{если } b > 0. \end{cases}$$

Определение 2. *Две точки $\overline{x}, \overline{y} \in \mathbb{R}$ эквивалентны относительно гиперплоскостей $l_1, .., l_k$ тогда и только тогда, когда $\sigma(\overline{x}) = \sigma(\overline{y})$, обозначим это через $\overline{x} \sim \overline{y}$.*

Легко проверить, что отношение "$\sim$" является отношением эквивалентности. Таким образом, пространство $\mathbb{R}^n$ разбивается на классы эквивалентности $R_1, .., R_s$.

Определение 3. *Сигнатурой класса R называется вектор $\sigma(R) = \sigma(\overline{x})$, где $\overline{x}$ точка класса R.*

Пусть $R_1, .., R_s$ - все классы эквивалентности на которые гиперплоскости $l_1, .., l_k$ разбивают $\mathbb{R}^n$.

Определение 4. *Функция $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ называется кусочно-линейной, если $\forall j \in \{1, .., s\}$ найдутся $b_j \in \mathbb{R}^n$ и $d_j \in \mathbb{R}$, что для всех $\overline{x} \in R_j$ выполняется $f(\overline{x}) = \overline{x} * \overline{b_j} + d_j$. Линейную функцию $\overline{x} * \overline{b_j} + d_j$, реализуемую на множестве R_j, обозначим $f_{R_j}(\overline{x})$. Множество всех кусочно-линейных функций обозначим через PL.*

Определение 5. *Функция $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ называется кусочно-постоянной, если $\forall j \in \{1, .., s\}$ найдутся $d_j \in \mathbb{R}$, что для всех $\overline{x} \in R_j$ выполняется $f(\overline{x}) = d_j$. Множество всех кусочно-постоянных функций обозначим через PC.*

Определение 6. *Функция $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ называется кусочно-линейной непрерывной, если $f \in PL$ и непрерывна. Класс всех кусочно-линейных непрерывных функций обозначим через CPL.*

Пусть $f : \mathbb{R}^n \Rightarrow \mathbb{R} \in CPL$, тогда выполнятся следующее свойство.

Свойство 1. *$\forall R_i, R_j \in \{R_1, .., R_s\}$, $\sigma(R_i) = (\sigma_1^{R_i}, .., \sigma_k^{R_i})$, $\sigma(R_j) = = (\sigma_1^{R_j}, .., \sigma_k^{R_j})$ таких, что*

$$\sum_{p=1}^{k} |\sigma_p^{R_i} - \sigma_p^{R_j}| = 1, \tag{1}$$

имеем

$$f_{R_i}(\overline{x}) = f_{R_j}(\overline{x}), \forall \overline{x} \in R_j. \tag{2}$$

В данном случае вектор сигнатуры R_j имеет на один ноль больше, иначе мы бы рассматривали $\overline{x} \in R_i$.

Определение 7. *Функция $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ называется согласованной, если $f \in PL$ и $\forall \overline{a}, \overline{b}, \overline{d} \in \mathbb{R}^n, \exists A, B, N \in \mathbb{R}$ такие, что $f(\overline{a} \cdot t + \overline{b}) = f(\overline{a} \cdot t + \overline{d}) + A \cdot h + B \cdot (h - 1)$, где $h = 0$ если $t < -N, h = 1$ если $t > N$ [4]. Класс всех согласованных функций обозначим P.*

Другими словами, функция f принадлежит классу согласованных функций, если на любых двух параллельных прямых в $\mathbb{R}^n$ функция определенная как разность функций на этих прямых, при достаточно большом $|t| > N$, ведет себя как кусочно-постоянная функция.

Задача вложености класса CPL в класс P.

Теорема 1. *$CPL \subset P$.*

Доказательство.

Докажем что любая функция из класса CPL принадлежит классу P.
От противного.
Пусть $f \in CPL$ и пусть $\exists \overline{a}, \overline{b}, \overline{d} \in \mathbb{R}^n$ такие, что $\forall N \in \mathbb{R}_+, f(\overline{a} \cdot t + \overline{b}) - f(\overline{a} \cdot t + \overline{d}) \neq const$, при $t > N$. Прямые $r_1 = \overline{a} \cdot t + \overline{b}$ и $r_2 = \overline{a} \cdot t + \overline{d}$ задают плоскость p в $\mathbb{R}^n$. Рассмотрим функцию f на плоскости p и обозначим ее через $g(x, y)$.

$$g(x, y) = f(p) \tag{3}$$

Далее будем рассматривать функции при достаточно большом значении аргумента таком, что область определения функции будет попадать только в один класс эквивалентности.
Пусть функция $g(x, y)$ задается гиперплоскостями (прямыми) $l_1, .., l_k$, которые образуют классы эквивалентности $R_1, .., R_s$.
Рассмотрим случаи:
1) $r_1, r_2 \subset R_i$.

$$g(\overline{a}\cdot t+\overline{b})-g(\overline{a}\cdot t+\overline{d}) = g_{R_i}(\overline{a}\cdot t+\overline{b})-g_{R_i}(\overline{a}\cdot t+\overline{d}) = A\cdot(a_1\cdot t+a_2\cdot t+b)+ \\ + B - A \cdot (a_1 \cdot t + a_2 \cdot t + d) + B = A \cdot (b - d) = const. \tag{4}$$

2) $r_1 \subset R_i, r_2 \subset R_j$, где $R_i \subseteq l \in \{l_1, .., l_k\}, R_j-$ смежен с прямой $l \in l_1, .., l_k$
Это означает что

$$\sum_1^s |\sigma^{R_i} - \sigma^{R_j}| = 1. \tag{5}$$

а следовательно $g_{R_i}(\overline{x}) = g_{R_j}(\overline{x}), \forall \overline{x} \in R_i$. Тогда

$$g(\overline{a}\cdot t+\overline{b})-g(\overline{a}\cdot t+\overline{d}) = g_{R_i}(\overline{a}\cdot t+\overline{b})-g_{R_i}(\overline{a}\cdot t+\overline{d}) = A\cdot(a_1\cdot t+a_2\cdot t+b)+ \\ + B - A\cdot(a_1\cdot t + a_2\cdot t + d) + B = A\cdot(b-d) = const. \quad (6)$$

3) Последний случай, когда между прямыми r_1 и r_2 проходят несколько прямых $l_{j_1}, .., l_{j_m} \in \{l_1, .., l_s\}$ упорядоченных от r_1 к r_2. Из первого и второго случаев можно показать что функция на прямой r_1 отличается на константу от функции на прямой l_{j_1}, функция на прямой l_{j_1} отличается на константу от функции на прямой l_{j_2}, и т.д. Получаем что функция на прямой r_1 отличается на константу от функции на прямой r_2.

Во всех трех случаях получаем что разница $g(\overline{a}\cdot t+\overline{b})-g(\overline{a}\cdot t+\overline{d}) = const$, а следовательно предположение неверно и класс $CPL \subset P$. Неравенство класса CPL и P вытекает из следующего примера. Рассмотрим функцию Хэвисайда $\Theta(x)$.

$$\Theta(x) = \begin{cases} 1, & \text{если } x \geq 0 \\ 0, & \text{если } x < 0 \end{cases}$$

В точке $x = 0$ функция разрывна, следовательно $\Theta(x) \notin CPL$. Но разность функции на двух параллельных прямых, при достаточно большом t, равно нулю.

$$\Theta(a\cdot t + b) - \Theta(a\cdot t + d) = 0. \quad (7)$$

Следовательно $\Theta(x) \in P$.
■

Класс CFP финитно-параллельных функций.

Замкнутость класса CFP.

Определение 8. *Функция $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ называется финитно-параллель- ной непрерывной функцией, если $f \in CPL$ и $\forall \overline{a}, \overline{b} \in \mathbb{R}^n, \exists N \in \mathbb{R}$ такое, что $f(\overline{a}\cdot t+\overline{b}) + f(\overline{a}\cdot(-t)+\overline{b}) = const$, для $t > N$. Множество всех финитно-параллельных непрерывных* *функций обозначим через CFP.*

Теорема 2. *Класс CFP замкнут по операциям суперпозиции.*

Доказательство.

Очевидно, что операции отождествления, переименования, добавления и удаления фиктивных переменных сохраняют класс CFP. Это вытекает из определения класса CFP. Докажем что операция подстановки не выводит нас за пределы класса CFP. Рассмотрим функцию $h(x_1, .., x_n)$.

$$h(x_1, ..., x_n) = f(g(x_1, .., x_n), x_2, .., x_n) \tag{8}$$

$f \in CFP, g \in CFP$.
Рассмотрим функцию h на произвольной прямой и обозначим через h_1:

$$\begin{aligned} h_1(t) = h(a_1 \cdot t + b_1, .., a_n \cdot t + b_n) = \\ = f(g(a_1 \cdot t + b_1, .., a_n \cdot t + b_n), a_2 \cdot t + b_2, .., a_n \cdot t + b_n) \end{aligned} \tag{9}$$

где $a_1, .., a_n, b_1, .., b_n \in \mathbb{R}$.
Так как $g \in CFP$ то $\exists N \in \mathbb{R}$ такое, что при $t > N$

$$\begin{aligned} h_1(t) = h(a_1 \cdot t + b_1, .., a_n \cdot t + b_n) = \\ = f(g(a_1 \cdot t + b_1, .., a_n \cdot t + b_n), a_2 \cdot t + b_2, .., a_n \cdot t + b_n) = \\ = f(A \cdot t + B, a_2 \cdot t + b_2, .., a_n \cdot t + b_n) \end{aligned} \tag{10}$$

При $t < -N$

$$\begin{aligned} h_1(t) = h(a_1 \cdot t + b_1, .., a_n \cdot t + b_n) = \\ = f(g(a_1 \cdot t + b_1, .., a_n \cdot t + b_n), a_2 \cdot t + b_2, .., a_n \cdot t + b_n) = \\ = f(A \cdot t + D, a_2 \cdot t + b_2, .., a_n \cdot t + b_n) \end{aligned} \tag{11}$$

Так как $f \in CFP \subset P \Rightarrow h_1(t) + h_1(-t) = const.$
■

Определение 9. *Класс A 2-предполный в классе B, если $B^{(2)} \subseteq [A \cup \{f\}]$, где $f \notin A$.*

Из теоремы 2 получили, что CFP замкнут в классе CPL. Оказывается данный класс является 2-предполным в классе CPL кусочно-линейных непрерывных функций.

2-предполнота класса CFP в классе CPL.

Теорема 3. *Пусть* $M \subseteq CPL$. $|x| \in [M \cup L] \Leftrightarrow M \not\subseteq CFP$.

Доказательство.

Необходимость вытекает из замкнутости класса CFP.

Докажем что из условия $M \not\subseteq CFP$ следует, что $|x| \in [M \cup L]$.
Пусть $M \not\subseteq CFP \Rightarrow \exists f(x_1, .., x_n)$ такая, что $\exists \overline{a}, \overline{b} \in \mathbb{R}^n$,
$f(\overline{a} \cdot t + \overline{b}) - f(\overline{a} \cdot (-t) + \overline{b}) \neq const$, $\forall N \in \mathbb{R}$, при $t > N$.
Пусть $g(t) = f(a_1 \cdot t + b_1, .., a_n \cdot t + b_n)$.

$$g(t) = \begin{cases} a_1 \cdot t + b_1, & \text{если } t \leq c_1 \\ a_2 \cdot t + b_2, & \text{если } c_1 \leq t \leq c_2 \\ \quad\cdot & \\ \quad\cdot & \\ a_l \cdot t + b_l, & \text{если } t \geq c_{l-1} \end{cases}$$

Сначала определим функцию $g_1(t)$.

$$g_1(t) = g(t + c_1) - a_2 \cdot t \tag{12}$$

Мы просто сместили функцию так, чтобы первый изгиб(слева направо) был в точке ноль. При этом отняли линейную функцию $a_2 \cdot t$, чтобы функция $g_1(t)$ на отрезке $[0; c_2 - c_1]$ равнялась нулю.
Далее определим функцию $h(t)$.

$$h(t) = \frac{(g(t + c_1) - a_l \cdot t)}{a_1 - a_l} \tag{13}$$

$a_1 - a_l \neq 0$ (следует из $h(t) \notin CFP$).
Данная функция также имеет первый изгиб в точке ноль. Отняв функцию $a_l \cdot t$ получили что на полуинтервале $[c_l - c_1; \infty)$ функция $h(t) = const$, при этом на полуинтервале $(\infty, 0], h(t) = t$.
Так, как функция $h(t)$ на полуинтервале $[c_l - c_1; \infty)$ равна константе следует, что функция $h(t)$ ограничена на полуинтервале

$[0;\infty)$ и имеет максимум и минимум.

$$max_h = \underset{t\geq 0}{max}\{(h(t)\} \tag{14}$$

$$min_h = \underset{t\geq 0}{min}\{(h(t)\} \tag{15}$$

Если $min_h = max_h$, положим $min_h = 0$, а $max_h = 1$.
Пусть

$$w = \frac{c_2 - c_1}{max_h - min_h}$$

Определим функцию $h_1(t)$.

$$h_1(t) = h(t + min_h) - min_h \tag{16}$$

$$h_1(t) = \begin{cases} t, & \text{если } t \leq 0 \\ c(t), & \text{если } t > 0, \text{где } 0 \leq c(t) \leq max_h - min_h \end{cases}$$

Подставим функцию $h_1(t)$ в функцию $g_1(t)$.

$$h_2(t) = g_1(w \cdot h_1(t)) \tag{17}$$

$$h_2(t) = \begin{cases} A \cdot t, & \text{если } t \leq 0 \\ 0, & \text{если } t \geq 0, \end{cases}$$

Из полученной функции $h_2(t)$ легко получить функцию $|t|$.

$$|t| = \frac{2}{A} \cdot (h_2(t) - \frac{A}{2} \cdot t) \tag{18}$$

■

Следствие 1. *Из теоремы 3 следует, что функию $|x|$ можно получить из произвольной функции не принадлежащей классу CFP. Причем нелинейная глубина и сложность равна двум.*

Мы выделяем функцию $|x|$ так, как она является удобной для построения произвольной двуместной кусочно-линейной непрерывной функции.
Обозначим через $CPL^{(1)}$ - все кусочно-линейные непрерывные функции зависящие от одной переменной.

Теорема 4. *Пусть* $M \subseteq CPL$. $CPL^{(1)} \subseteq [M \cup L] \Leftrightarrow M \not\subseteq CFP$.

Данная теорема была сформулированна и доказана в работе [3]. Из теоремы следует что все одноместные кусочно-линейные непрерывные функции выражаются через модуль и линейные функции. Следующая теорема утверждает что все двуместные кусочно-линейные непрерывные функции также выражаюется через модуль и линейные функции.

Теорема 5. *Пусть* $M \subseteq CPL$. $CPL^{(2)} \subseteq [M \cup L] \Leftrightarrow M \not\subseteq CFP$.

Доказательство.
Определим следующие функции.

$$min(x, y) = \frac{x + y - |x - y|}{2} \tag{19}$$

$$max(x, y) = \frac{x + y + |x - y|}{2} \tag{20}$$

$MM = \{min, max\} \subset CPL$.

Пусть $f \in CPL^{(2)}$ задается гиперплоскостями (прямыми) $l_1, .., l_k$. Данные прямые разбивают плоскость $\mathbb{R}^2$ на классы эквивалентности $R_1, .., R_s$. Выберем произвольную прямую l_i и рассмотрим все тройки
(R_+, R_0, R_-), где $R_+, R_0, R_- \in \{R_1, .., R_s\}$ смежные с прямой l_i и таких, что вектор сигнатуры для R_+ не имел нуля [1]. Без ограничения общности можно считать что $f_{R_+} = f_{R_0}, \forall(R_+, R_0, R_-)$ (следует из определения класса CPL). Прономеруем все тройки $(R_+, R_0, R_-)_j, j = 1..w$ "слева направо"относительно прямой l_i. Рассмотрим первую тройку $(R_+, R_0, R_-)_1$. Положим

$$l_{-;i,1} = f_{R_-} - f_{R_0} \tag{21}$$

$$f_{i,1} = f - m(0, l_{-;i,1}) \tag{22}$$

где $m \in MM$. если $l_{-;i,1} > 0$ на R_-, то $m = max$. В противном случае $m = min$.

Так как функция $m(0, l_{-;i,1})$ меняет значение только в нижней части функции f относительно прямой l_i, то для тройки $(R_+, R_0, R_-)_1$, выполняется условие

$$f_{R_+} - f_{R_0} = 0 \text{ и } f_{R_-} - f_{R_0} = 0. \tag{23}$$

Функция $m(0, l_{-;i,1})$ не добавляет новых задающих прямых к уже имеющимся.

Рассмотрим следующую тройку $(R_+, R_0, R_-)_2$. Если данная тройка отделена от предыдущей только одной прямой, то нетрудно видеть, что она тоже удовлетворяет условию (23). Если между текущей и предыдущей тройкой существует несколько разделяющих прямых, то существует класс эквивалентности R_c, который имеет общую границу с классом R_- и имеет одну общую точку с прямой l_i. Тогда положим:

$$l_{-;i,2} = f_{R_-} - f_{R_0} \tag{24}$$

$$l_{R_c;i,2} = f_{R_c} - f_{R_0} \tag{25}$$

$$f_{i,2} = f_{i,1} - b_{i,2} \tag{26}$$

где $b_{i,2} = m_1(0, m_2(l_{-;i,2}, l_{R_c;i,2})), m_1, m_2 \in MM$. Если $l_{-;i,1} > 0$ на R_-, то $m_1 = max, m_2 = min$. В противном случае $m_1 = min, m_2 = max$.

Функция $b_{i,2}$ меняет значение только в нижней части функции f относительно прямой l_i. При этом не меняет значение функции на предыдущих тройках (R_+, R_0, R_-). Функция $b_{i,2}$ не добавляет новых задающих прямых.Тогда для тройки $(R_+, R_0, R_-)_2, f_{R_+} - f_{R_0} = 0$ и $f_{R_-} - f_{R_0} = 0$. И так далее $\forall (R_+, R_0, R_-)_j, j = 1..w$ получим что $f_{R_+} - f_{R_0} = 0$ и $f_{R_-} - f_{R_0} = 0$, а следовательно переход через прямую l_i тривиален и функцию $f_{i,w}$ можно задать прямыми $l_1, .., l_{i-1}, l_{i+1}, .., l_k$.

Повторив эту процедуру для каждой прямой $l_1, .., l_k$ получим что конечная функция задается пустым множеством прямых. Это означает что переход через любую прямую тривиален и конечная функция линейна [1]. Следовательно любую функцию из $CPL^{(2)}$ можно получить из функции $|x|$ и линейных функций используя операции суперпозиции.

■

Следствие 2. *Из теоремы 5 следует, что произвольную двуместную кусочно-линейную непрерывную функцию можно получить из модуля и линейных функций по операциям суперпозиции. Причем нелинейная глубина не превышает четырех, а нелинейныя сложнать асимптотически равна* $O(n^2)$*, где n это количество задающих прямых.*

Список литературы

[1] Половников В.С. «Об оптимизации структурной реализации нейронных сетей». Диссертация на соискание учёной степени кандидата физико-математических наук. Москва 2007.

[2] Кудрявцев В.Б., Алешин С.В., Подколзин А.С. «Введение в теорию автоматов». Издательство «Наука», Москва, 1985 г. 1-320 стр.

[3] Самсонова Н.А. «Класс кусочно-линейных непрерывных функций». Бакалаврская работа. Якутск 2015.

[4] Кан А.Н. «Вопросы выразимости в классе нейронных функций с памятью». Бакалаврская работа. Ташкент 2014.