Вопросы выразимости в классе нейронных функций

А. Н. Кан

Рассматривается класс PL кусочно-линейных функций вместе с операциями суперпозиции [1]. В настоящей работе показано, что в PL существуют три предполных класса, содержащих класс L всех линейных функций: класс финитных функций, класс непрерывных функций и класс согласованных функций. Получен критерий позволяющий по конечному множеству $M \in PL$ проверить полноту множества $M \cup L$.

Ключевые слова: класс кусочно-линейных функций, класс финитных функций, класс непрерывных функций, класс согласованных функций, функция Хэвисайда, операции суперпозиции, существенный разврыв.

Рассмотрим следующие функции действительных аргументов:

1. Линейные функции: $c_1x_1 + c_2x_2 + \ldots + c_nx_n + c_0$, $c_i \in \mathbb{R}$, $n \in \mathbb{N}$, $i = 0, 1, \ldots, n$.

2. Функция Хэвисайда: $\Theta(x) = \begin{cases} 1, & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0. \end{cases}$

3. Функция $F(x, y) = \begin{cases} x, & \text{если } y \geqslant 0, \\ 0, & \text{если } y < 0. \end{cases}$

Следущие определение и утверждение содержатся в работе [1].

Определение. Кусочно линейная функция называется финитной если она принадлежит следующему классу $\Phi = \{f(x_1, \ldots, x_n) : f \in PL$, $\forall a_i, b_i \in \mathbb{R}$, $i = 1 \ldots n$, $\exists c, d, N \in \mathbb{R}$, при $\mid t \mid > N$, $f(a_1 * t + b_1, \ldots, a_n * t + b_n) = c * t + d\}$.

Утверждение 1. *Класс Φ замкнут по операциям суперпозиции.*

Множество, состоящее из всех кучочно-линейных непрерывных функций, обозначим через C.

Утверждение 2. *Класс C замкнут по операциям суперпозиции.*

Лемма 1. *Пусть M некоторое подмножество кусочно-линейных функций. Если $M \nsubseteq C$ и $M \nsubseteq \Phi \Rightarrow \exists f \in [M \cup L]$ такая, что $f(a_1 * t + b_1, \dots, a_n * t + b_n)$ имеет существенный разрыв.*

Теорема 1. *Пусть M некоторое подмножество кусочно-линейных функций. $\Theta(x) \in [M \cup L] \Leftrightarrow M \nsubseteq C$ и $M \nsubseteq \Phi$.*

Доказательство. Покажем, что из $M \nsubseteq C$ и $M \nsubseteq \Phi \Rightarrow \Theta(x) \in [M \cup L]$.

Пусть функция $g \notin \Phi \Rightarrow \exists a_i, b_i, N$ такие, что $g(a_1 * t + b_1, \dots, a_n * t + b_n) = A * t + B$, для $t > N$ и $g(a_1 * t + b_1, \dots, a_n * t + b_n) = C * t + D$, для $t < -N$, где $A, B, C, D \in \mathbb{R}$, $((A-C)^2 + (B-D)^2) \neq 0$.

Положим $G(t) = g(a_1 * (t+h) + b_1, \dots, a_n * (t+h) + b_n) - g(a_1 * (-t) + b_1, \dots, a_n * (-t) + b_n)$:

$$G(t) = \begin{cases} (A+C) * t + (B-D) + A * h, \text{ при } t > N + h, \\ (A+C) * t + (D-B) + C * h, \text{ при } t < -(N+h). \end{cases}$$

Если $B - D = D - B = 0$, то $h = 1$, иначе $h = 0$; Имеем, что функция $G \notin \Phi$.

Положим

$$G_1(t) = \left(\frac{G(t) - ((A+C) * t + (D-B) + C * h)}{2 * (B-D) + (A-C) * h} \right);$$

$$G_1(t) = \begin{cases} 1, \text{ при } t > N + h, \\ 0, \text{ при } t < -(N+h). \end{cases}$$

По лемме 1 $\exists f \in [M \cup L]$ такая, что $f(a_1 * x + b_1, \dots, a_n * x + b_n)$ имеет существенный разрыв в точке $x = T$. Сделаем замену $t = (x - T)$, тогда $p(t) = f(a_1 * (t+T) + b_1, \dots, a_n * (t+T) + b_n)$ имеет существенный разрыв в точке $t = 0$. Будем считать, что в некоторой ϵ-окрестности точки $t = 0$, значение функции $p(t)$ слева всюду меньше чем справа. Это, при необходимости, можно достичь заменой $z = -t$.

Положим $F(t) = p(t) - A + B * t$, где

$$A = \begin{cases} (p(-0) + p(+0))/2, \text{ при } p(0) - (p(-0) + p(+0))/2 \neq 0 \\ p(-0) + p(+0))/2 + r, 0 < r < (p(-0) + p(+0))/2, \\ \qquad \text{при } p(0) - (p(-0) + p(+0))/2 = 0. \end{cases}$$

$B = \max(|(p(t) - A)/t|) + 1, t \neq 0$, тогда $F(t) < 0$, при $t < 0$; $F(t) > 0$, при $t > 0$.

$$\Theta(x) = \begin{cases} G_1((N+1+h)*(F(x)/\min(|F|))), & \text{при } F(0) > 0, \\ -G_1((N+1+h)*(F(-x)/\min(|F|))) + 1, & \text{при } F(0) < 0. \end{cases}$$

Теорема доказана.

Определение. Кусочно-линейная функция называется согласованной если она принадлежит следующему классу $P = \{f(x_1, \ldots, x_n) : f \in PL, \forall a_i, b_i, d_i \in \mathbb{R}, i = 1 \ldots n, \exists A, B, N \in \mathbb{R}$, при $|t| > N, f(a_1 * t + b_1, \ldots, a_n * t + b_n) = f(a_1 * t + d_1, \ldots, a_n * t + d_n) + A * h + B * (h-1)$, где $h = 0$, если $t < -N$, $h = 1$, если $t > N\}$.

Утверждение 3. *Класс P замкнут по операциям суперпозиции.*

Теорема 2. *Пусть M некоторое подмножество кусочно-линейных функций. $F(x,y) \in [M \cup L] \Leftrightarrow M \not\subseteq P, M \not\subseteq C$ и $M \not\subseteq \Phi$.*

Доказательство. Покажем, что из $M \not\subseteq P, M \not\subseteq C$ и $M \not\subseteq \Phi \Rightarrow F(x,y) \in [M \cup L]$. Из теоремы 1 $\Rightarrow \Theta(x) \in [M \cup L]$.

Пусть функция $g \notin P \Rightarrow \exists a_i, b_i, d_i, A_1, A_2, B_1, B_2, C_1, C_2, D_1, D_2$ и $N \in \mathbb{R}$ такие, что $g(a_1 * t + b_1, \ldots, a_n * t + b_1) = A_1 * t + B_1$, при $t > N$ и $g(a_1 * t + b_1, \ldots, a_n * t + b_n) = A_2 * t + B_2$, при $t < -N$. $g(a_1 * t + d_1, \ldots, a_n * t + d_1) = C_1 * t + D_1$, при $t > N$ и $g(a_1 * t + d_1, \ldots, a_n * t + d_n) = C_2 * t + D_2$, при $t < -N$. Где $A_1 \neq C_1$.

Положим $w_i = d_i - b_i, i = 1 \ldots n$. Построим функцию $g_1(t,y) = g(a_1 * t + b_1 + w_1 * \Theta(\Theta(t) + \Theta(y) - 2), \ldots, a_n * t + b_n + w_n * \Theta(\Theta(t) + \Theta(y) - 2))$; тогда $g_1(t,y) = \begin{cases} g(a_1 * t + d_1, \ldots, a_n * t + d_1), & \text{при } t \geqslant 0, y \geqslant 0 \\ g(a_1 * t + b_1, \ldots, a_n * t + b_1), & \text{иначе.} \end{cases}$. Далее отнимем от функции $g_1(t,y)$ функцию $g_1(-t,y)$. Полученная функция будет равняться одной и той же линейной функции независимо от параметра t (при достаточно больших t), но в зависимости от параметра y будет равняться разным линейным функциям. Заметим, что проделав аналогичные операции с функцией g, мы бы необязательно получили функцию обладающей такими свойствами.

Положим $G(t,y) = g_1(t,y) - g_1(-t,y)$, тогда

$$G(t,y) = \begin{cases} (A_1 + A_2) * t + (B_1 - B_2), & \text{при } |t| > N, y < 0, \\ (C_1 + A_2) * t + (D_1 - B_2), & \text{при } |t| > N, y \geqslant 0. \end{cases}$$

Далее избавимся от констант $B_1 - B_2$ и $D_1 - B_2$. Положим $G_1(t,y) = G(t,y) - (B_1 - B_2) * (1 - \Theta(y)) - (D_1 - B_2) * \Theta(y)$ тогда $G_1(t,y) = \begin{cases} (A_1 + A_2) * t, \text{ при } |t| > N, y < 0, \\ (C_1 + A_2) * t, \text{ при } |t| > N, y \geqslant 0. \end{cases}$ Сделаем так чтобы функция равнялась константе 0, при $y < 0$. Положим $G_2(t,y) = G_1(t,y) - (A_1 + A_2) * t$, тогда

$$G_2(t,y) = \begin{cases} 0, \text{ при } |t| > N, y < 0, \\ (C_1 - A_1) * t, \text{ при } |t| > N, y \geqslant 0. \end{cases}$$

Положим $G_3(t,y) = (1/(C_1 - A_1)) * G_2(t,y)$;

$$G_3(t,y) = \begin{cases} 0, \text{ при } |t| > N, y < 0, \\ t, \text{ при } |t| > N, y \geqslant 0. \end{cases}$$

Функция $G_3(t,y)$ определена только для $|t| > N$. Пусть $t = x + (2 * (N+1) * \Theta(x) - (N+1))$; Положим $F(x,y) = G_3(t,y) - (2 * (N + 1) * \Theta(\Theta(x) + \Theta(y) - 2) - (N+1))$, тогда $F(x,y) = \begin{cases} 0, \text{ при } y < 0, \\ x, \text{ при } y \geqslant 0. \end{cases}$ Получили искомую функцию. Теорема доказана.

Список литературы

[1] Половников В. С. Об оптимизации структурной реализации нейронных сетей / Дисс. на соиск. уч. ст. к.ф.-м.н. — М., 2007.

[2] Кудрявцев В. Б., Алешин С. В., Подколзин А. С. Введение в теорию автоматов. — М.: Наука, 1985.