О классе нейронных функций с двоично-рациональными параметрами

Агафонова М.В.

В работе рассматривается класс нейронных кусочно-параллельных функций с двоично-рациональными коэффициентами. Доказано что для любой функции из класса кусочно-параллельных функций существует функция из класса кусочно-параллельных функций с двоично-рациональными коэффициентами, приближающая ее с наперед заданной точностью. Также, для рассматриваемого класса функций показано, что существуют базисы, состоящие из заданного (произвольного) числа элементов, в частности, найдена Шефферова функция.

Ключевые слова: класс кусочно-параллельных функций, класс нейронных функций с двоично-рациональными коэффициентами, операции суперпозиции, Шефферова функция, базисы.

1. Введение.

В данной работе рассматривается класс кусочно-параллельных функций с двоично-рациональными коэффициентами (BPP), являющийся аппроксимирующим для рассмотренного в работе [1] класса кусочно-параллельных функций (PP). Основной темой исследования работы [1] является изучение, так называемых нейронных схем. Под понятием нейронная схема понимается математический объект, представляющий из себя набор функциональных элементов, определенных строением модели исследуемой нейронной сети, а также функциональные схемы, полученные из них операциями суперпозиции. Где под суперпозицией понимаются операции: добавления фиктивного входа, изъятия фиктивного входа, склеивания входов, переименования входов без склеивания, последовательного соединения [1]. Стоит отметить, что данное построение происходит по аналогии с построением автоматных схем [2].

Особый же интерес в работе [1] представляет доказательство совпадения множеств функций, реализуемых нейронными схемами, с некоторыми классами кусочно-линейных функций, изучение которых в дальнейшем и ведется. Так, например, устанавливается эквивалентность между множеством функций, реализуемых нейронными схемами без памяти, и множеством кусочно-линейных функций (PL), эквивалентность множества функций, реализуемых нейронными схемами модели Мак-Каллока-Питтса[3], множеству кусочно-параллельных функций (PP).

Изучение класса кусочно-линейных функций (PL) получило развитие в работе [4]. В настоящей же работе, находит свое развитие изучение подкласса кусочно-параллельных функций PP. Рассматриваемый класс, являющийся замыканием [5] множества B, $B = \{\frac{1}{2}x, -x, x+y, \theta(x)\}$, будем называть классом кусочно-параллельных функций с двоично-рациональными коэффициентами, и обозначать BPP. Интерес изучения класса BPP состоит в том, что он является конечно-порожденным, в отличие от ранее рассматриваемых классов в работах[1], [4]. В настоящей работе показано, что для любой функции из класса PP существует функция из класса BPP, приближающая ее с наперед заданной точностью. Далее, абстрагируясь от смысловой интерпретации этих функций, рассматриваются задачи типичные для всех функциональных систем [6]. Устанавливается существование Шефферовой функции[5]. Такой функцией является: $F(x,y,z) = x - \frac{1}{2}y - \frac{1}{2}\theta'(z) + \frac{1}{2}$, где $\theta'(z) = \begin{cases} 1 & \text{при } z > 0 \\ 0 & \text{при } z \leq 0 \end{cases}$.

Доказано существования в классе BPP базиса из k элементов для произвольного натурального числа k.

2. Основные определения.

Определение1. Множество кусочно-параллельных функций, определяется следующим выражением:

$PP = \{f | f = f_c + f_l, f_c \in PC, f_l \in L\}$, где PC – множество всех кусочно-постоянных функций, L – множество всех линейных функций [1]

Рассмотрим множество функций $B = \{\frac{1}{2}x, -x, x+y, \theta(x)\}$, в котором $x \in \mathbb{R}, y \in \mathbb{R}$. В множестве B содержатся функции: умножитель на двоично-рациональную константу $f_1(x) = \frac{1}{2}x$, отрицание $f_2(x) = -x$, сумматор $f_3(x,y) = x + y$, функция Хэвисайда $\theta(x) = \begin{cases} 1 & \text{при } x \geq 0 \\ 0 & \text{при } x < 0 \end{cases}$.

Определение 2. Замыкание функций, принадлежащих множеству $B = \{\frac{1}{2}x, -x, x+y, \theta(x)\}$, по операциям суперпозиции [1],[4] назовем множеством кусочно-параллельных функций с двоично-рациональными коэффициентами и обозначим BPP.

$$[B] = BPP.$$

Данное замыкание содержит все целые и двоично-рациональные константы. Так константа $\frac{1}{2}$ может быть получена подстановкой: $\frac{1}{2}(\theta(\theta(x))$. Константа 1 получается из f_3 подстановкой: $f_3(\frac{1}{2}, \frac{1}{2}) = 1$. Все положительные целые числа можно получить также из $f_3 : f_3(1, n-1) = 1 + (n-1)$. Произведя затем, подстановки вида:$f_2(n) = -n$, получим все целочисленные константы. Двоично-рациональные константы могут быть получены при помощи подстановок функций $f_1(\frac{1}{2^{k-1}}) = \frac{1}{2}(\frac{1}{2^{k-1}}) = \frac{1}{2^k}$ и $f_3(\frac{1}{2^k}, \frac{m-1}{2^k}) = \frac{m}{2^k}$.

3. Теорема о приближении с наперед заданной точностью функций, принадлежащих классу PP, функциями класса BPP.

Определение 3. Пусть $g(x_1, x_2, \ldots, x_n) \in$ PP, $f(x_1, x_2, \ldots, x_n) \in$ BPP. Функция $f(x_1, x_2, \ldots, x_n)$ приближает функцию $g(x_1, x_2, \ldots, x_n)$ на множестве $A = [0,1]^n$ с точностью ε,если $\forall\ n\ \exists\ A' \subseteq [0,1]^n$ такое что $|A'| > 1 - \varepsilon$ и $\forall (x_1, x_2, \ldots, x_n) \in A' : |f(x_1, x_2, \ldots, x_n) - g(x_1, x_2, \ldots, x_n)| < \varepsilon$.

Теорема 1. *$\forall g(x_1, x_2, \ldots, x_n) \in PP\ \forall\ \varepsilon\ \exists\ f(x_1, x_2, \ldots, x_n) \in BPP$, которая приближает g на множестве $A = [0,1]^n$ с точностью ε.*

Доказательство: Сначала определим $A' \subseteq [0,1]^n$.
Функцией $g(x_1, x_2, \ldots, x_n) \in$PP, порождается k гиперплоскостей $l_1, \ldots, l_k$ разбивающих R^n на классы эквивалентности $R_1, \ldots, R_s, (k, s \in \mathbb{N})$. Обозначим через $k'(k' \in 1, \ldots, k)$ – количество гиперплоскостей $l_1, \ldots, l_{k'}$, пересекающих рассматриваемый единичный гиперкуб A.

При $k' = 0, \forall n$, положим $A' = A$. Построим множество A' при $k' = 1$, и затем обобщим для любого $k' \in \{1, \ldots, k\}$. Рассмотрим случаи $n = 1, n = 2$ и потом индуктивно построим A' для произвольного n.

Случай $k' = 1, n = 1$. То есть множество A представляет собой отрезок $[0,1]$, имеется одна разделяющая гиперплоскость, допустим проходящая через точку $B \in [0,1], B \in \mathbb{R}$. Если B лежит внутри интервала $(0,1)$, то приблизим B двоично-рациональными числами B' и B'' слева и с права соответственно, такими, что $|B - B'| \leq \frac{\delta}{2}, |B - B''| \leq \frac{\delta}{2}$. (рис. 1). Тогда

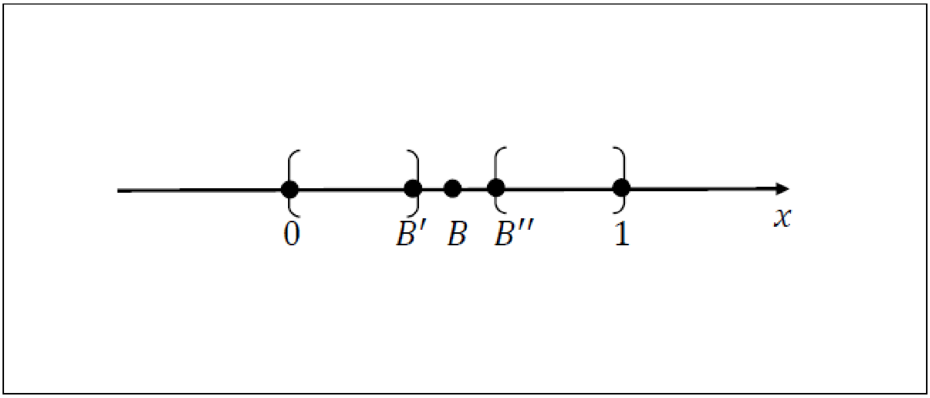

Рис. 1.

в качестве $A^{'}$ рассмотрим объединение множеств $[0, B^{'}] \cup [B^{''}, 1]$. Заметим, что для любого наперед заданного $\varepsilon > 0$ при $\delta < \varepsilon$, $|A^{'}| > 1 - \varepsilon$. Следовательно положим $A^{'} = [0, B^{'}] \cup [B^{''}, 1]$. Если точка B лежит на одном из концов отрезка $[0, 1]$, т.е. ее координатами будут целые числа либо 0, либо 1 соответственно. Поскольку, классу BPP принадлежит множество всех целочисленных констант, можно принять $B^{'} = 0 (B^{'} = 1), B^{''} = 0 (B^{''} = 1)$ и так как $|B - B^{'}| = 0, |B - B^{''}| = 0$, положить $A^{'} = A$. И следовательно свести этот случай к случаю $k = 0$. Стоит отметить, что так же можно поступить и когда B принимает любые двоично-рациональные координаты внутри отрезка $[0, 1]$.

Случай $k^{'} = 1, n = 2$, представляет больший интерес. Множество A в этом случае – это квадрат с единичной стороной. Разделяющая множество A гиперплоскость l, на A представляет отрезок прямой, лежащий внутри этого квадрата. Обозначим его BC. (рис.2). Случай, если отрезок лежит на одной из сторон, аналогичен предыдущему случаю, когда точка B совпадала с одним из концов отрезка. Диагональ квадрата обозначим OD. Длина отрезка $BC \leq OD$, в зависимости от его расположения (равенство достигается при совпадении BC с OD). Приблизим точки $B, C \in \mathbb{R}$ точками $B^{'}, B^{''}$ и $C^{'}, C^{''}$ соответственно. Причем $B^{'}, B^{''}, C^{'}, C^{''} \in \{\frac{m}{2^k}, m, k \in \mathbb{N}\}$, $|B - B^{'}| \leq \frac{\delta}{2}, |B - B^{''}| \leq \frac{\delta}{2}, |C - C^{'}| \leq \frac{\delta}{2}, |C - C^{''}| \leq \frac{\delta}{2}$. В зависимости от расположения BC, получим трапецию $B^{'}B^{''}C^{'}C^{''}$ с высотой $B^{'}E < B^{'}B^{''} = \delta$ и основаниями $B^{'}C^{'} = B^{'}C^{''} < OD$ или прямоугольник со сторонами $B^{'}B^{''} = C^{'}C^{''} = \delta$ и $B^{'}C^{'} = B^{'}C^{''} = BC$. Диагональ OD также достроим до прямоугольника $O^{'}O^{''}D^{'}D^{''}$ как показано на рисунке 2. $O^{'}O^{''} = D^{'}D^{''} = \delta, O^{'}D^{'} = O^{''}D^{''} = OD$.Заметим,

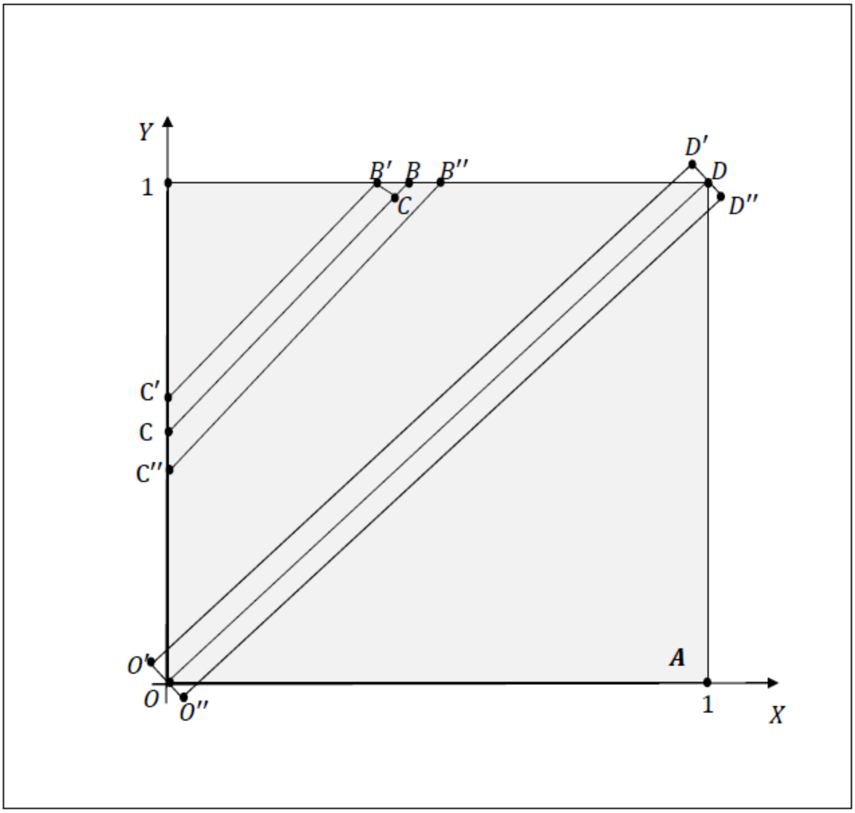

Рис. 2.

что $|O^{'}D^{'}| = \sqrt{2}$. Из соотношения BC$\leq$OD, следует, что

$$S_{B'B''C'C''} \leq S_{O'O''D'D''} = |O^{'}D^{'}| \cdot |O^{'}O^{''}| = \sqrt{2}\delta < \varepsilon.$$

Следовательно при $\delta < \frac{\varepsilon}{\sqrt{2}}$ множество $A^{'} = A \backslash S_{B'B''C'C''}, |A^{'}| > 1 - \varepsilon$.

Теперь обобщим этот случай для произвольного n (при $k' = 1$). Сечение n-мерного куба, будет иметь размерность $(n-1)$. Его пересечение с гранью n-мерного куба будет иметь размерность $(n-2)$. Гиперплоскость l будет проходить через одно из таких сечений. Рассмотрим внутри n – мерного куба A $(n-1)$-мерный многогранник с $(n-2)$-мерными сторонами представляющий это сечение соответствующее l. Причем каждая грань единичного n-мерного куба A представляет собой также единичный $(n-1)$-мерный куб. Например для $n = 3$ сечение, соответствующее l будет иметь вид многоугольника стороны которого являются отрезка-ми лежащими на гранях А, в свою очередь являющихся кубами меньшей размерности т.е. квадратами для $n = 3$. Для построения $A^{'}$ для n-мерного куба А, необходимо оценить объем Δ максимально возможного $(n-1)$-мерного сечения в А, затем приблизить его сечениями $\Delta^{'}, \Delta^{''}$ имеющими двоично рациональные координаты и такими, что:

$$|x_i - x_i^{'}| < \frac{\delta}{2}, |x_i - x_i^{''}| < \frac{\delta}{2}, \forall x_i \in \Delta, \forall x_i^{'} \in \Delta^{'}, \forall x_i^{''} \in \Delta^{''}.$$

Оценим объем Δ максимально возможного $(n-1)$-мерного сечения в n-мерном кубе А, отталкиваясь от идеи, что это будет $(n-1)$-мерный многогранник, со сторонами лежащими в $(n-1)$-мерных единичных кубах. Объём этого многогранника, будет меньше или равен объему $(n-1)$-мерного куба, с максимально возможными гранями лежащими в $(n-1)$-мерных кубах, являющихся гранями А. Максимально возможные грани в кубе можно оценить его диагональю равной $\sqrt{n}$.

$$S_{\Delta} \le (\sqrt{n})^{n-1}, n \ge 2.$$

Учитывая , приближения $\Delta^{'}, \Delta^{''}$, заметим, что $S_{\Delta^{'},\Delta^{''}}$ равна произведению высоты, в качестве которой мы возьмем δ, на основание, т.е. S_{Δ},следовательно получим оценку для$S_{\Delta^{'},\Delta^{''}}$:

$$S_{\Delta^{'},\Delta^{''}} \le (\sqrt{n})^{n-1}\delta < \varepsilon.$$

Следовательно, для $n \ge 2$, при $\delta < \frac{\varepsilon}{(\sqrt{n})^{n-1}}$, множество $A^{'} = A \backslash S_{\Delta^{'},\Delta^{''}}, |A'| > 1 - \varepsilon$.

Наконец обобщим результат для произвольного $k^{'}$. Тогда в качестве $A^{'}$ рассмотрим множество $A \backslash k^{'} S_{\Delta^{'},\Delta^{''}}$, тогда $|A^{'}| > |A| - k^{'}|S_{\Delta^{'},\Delta^{''}}| = 1 - k^{'}(\sqrt{n})^{n-1}) \cdot \delta$, и при $\delta < \frac{\varepsilon}{k \cdot (\sqrt{n})^{n-1})}, |A^{'}| > 1 - \varepsilon$. Отдельно стоит отметить, что в случае, $k^{'} > 1$ возможно, что концы отрезка для $n = 2$ и грани многогранника в случае $n > 2$ представляющего l, не будут лежать на гранях А. В этом случае продолжим многогранник по плоскости сечения соответствующего l (отрезок продолжим по прямой на которой он лежит) до пересечения с гранями А. Очевидно это только увеличит его объём и следовательно, найденная оценка сохранится.

Теперь построим такую функцию $f(x_1, x_2, \ldots, x_n) \in BPP$,что $\forall (x_1, x_2, \ldots, x_n) \in A^{'} : |f(x_1, x_2, \ldots, x_n) - g(x_1, x_2, \ldots, x_n)| < \varepsilon$. Пусть на некотором классе $R_i, i \in \{1, \ldots, s\}$, имеет место равенство:

$$g(x_1, x_2, \ldots, x_n) = c_n x_n + c_{n-1}x_{n-1} + \cdots + c_0, c_i \in \mathbb{R}.$$

Тогда рассмотрим пересечение: $R_i \cap A'$. Это множество задается пересечением гиперплоскостей. Определим на нем функцию:

$$f(x_1, x_2, \dots, x_n) = d_n x_n + d_{n-1} x_{n-1} + \dots + d_0, d_i \in \{\frac{m}{2^k}, m, k \in \mathbb{N}\}, 0 \leq x_i \leq 1,$$

приближающую $g(x_1, x_2, \dots, x_n)$. Пусть $|c_i - d_i| < \frac{\varepsilon}{n+1} = \delta', \forall i$. Так как $0 \leq x_i \leq 1$:

$$|c_i x_i - d_i x_i| \leq |c_i - d_i||x_i| \leq 1 \frac{\varepsilon}{n+1} \leq \frac{\varepsilon}{n+1}.$$

Следовательно

$$|c_n x_n + c_{n-1} x_{n-1} + \dots + c_0 - d_n x_n - d_{n-1} x_{n-1} - \dots - d_0| \leq$$

$$|c_n x_n - d_n x_n| + |c_{n-1} x_{n-1} - d_{n-1} x_{n-1}| + \dots + |c_0 - d_0| \leq$$
$$\frac{\varepsilon}{n+1} + \frac{\varepsilon}{n+1} \dots + \frac{\varepsilon}{n+1} \leq \varepsilon$$

Следовательно при $\delta' \leq \frac{\varepsilon}{n+1}$, выполняется неравенство:

$$|f(x_1, x_2, \dots, x_n) - g(x_1, x_2, \dots, x_n)| < \varepsilon$$

Теорема доказана.

4. Существование в классе функций ВРР базиса размерности равной любому наперед заданному натуральному числу k.

Сам класс ВРР, представляет множество кусочно-линейных функций от n переменных, с двоично-рациональными коэффициентами и возможными разрывами 1го рода. В соответствии с леммой о нелинейной глубине[1], все функции класса ВРР представляются в виде:

$$(*) f(x_1, x_2, \dots, x_n) = \frac{k_1}{2^{m_1}} x_1 + \frac{k_2}{2^{m_2}} x_2 + \dots + \frac{k_i}{2^{m_i}} x_i +$$

$$+ \frac{k_{i+1}}{2^{m_{i+1}}} \theta(\omega_1) + \dots + \frac{k_j}{2^{m_j}} \theta(\omega_s) + \frac{k_0}{2^{m_0}}$$

где, $0 \leq i \leq N, i \leq j \leq N$,
$k_1, k_2, \dots k_N \in \mathbb{Z}, (N \leq n)$,

$m_1, m_2, \dots m_N \in \mathbb{N}, (N \le n),$
$0 \le l \le s,$
$0 \le s \le n-i,$

$$\theta(\omega_l) = \theta(\frac{k'_{l0}}{2^{m'_0}} + \frac{k'_{l1}}{2^{m'_1}}x_1 + \frac{k'_{l2}}{2^{m'_2}}x_2 + \dots + \frac{k'_{li'}}{2^{m'_{i'}}}x_{i'} + \frac{k'_{li+1}}{2^{m'_{i+1}}}x^l_{i+1} + \frac{k'_{li+2}}{2^{m'_{i+2}}}x^l_{i+2} + \dots + \frac{k_{l'}}{2^{m_{j'-1}}}x^l_{j'-1} + \frac{k_{j'}}{2^{m_{j'}}}\theta(\omega_{l1}) + \dots + \frac{k_{t'}}{2^{m_{t'}}}\theta(\omega_{ls'}))$$

$$\theta(\omega_{li'}) = \theta(\frac{k'_{li'0}}{2^{m'_0}} + \frac{k'_{li'1}}{2^{m'_1}}x_1 + \frac{k'_{li'2}}{2^{m'_2}}x_2 + \dots + \frac{k'_{li'i}}{2^{m'_i}}x_i + \frac{k'_{li'i+1}}{2^{m'_{i+1}}}x^{li'}_{i+1} + \frac{k'_{li'i+2}}{2^{m'_{i+2}}}x^{li'}_{i+2} + \dots + \frac{k'_{li'j'-1}}{2^{m'_{j'-1}}}x_{i'})$$

$k'_{l0}, k'_{l1}, \dots, k'_t \in \mathbb{Z}, (N \le n)$
$k'_{li'0}, k'_{li'1}, \dots, k'_{li'j'-1} \in \mathbb{Z}, (N \le n)$
$m'_1, m'_2, \dots, , m'_t \in \mathbb{N}, (N \le n), 0 \le t \le s, 0 \le j' - 1 \le n, 0 \le i' \le i,$
$i + \sum_l (j' - 1 - i') + \sum_{l'} (j' - 1 - i') = n$

4.1. Шефферова функция для класса BPP.

Теорема 2. *В классе BPP, функция* $F(x, y, z) = x - \frac{1}{2}y - \frac{1}{2}\theta'(z) + \frac{1}{2}$, *где* $\theta'(z) = \begin{cases} 1 & \text{при } z > 0 \\ 0 & \text{при } z \le 0 \end{cases}$. *, является шефферовой.*

.

Доказательство: Сначала получим константу $\frac{1}{2}$. Для этого рассмотрим суперпозицию:

$F_1(x, y, z) = F(F(x, y, z), y, z) = x - \frac{1}{2}y - \frac{1}{2}\theta'(z) + \frac{1}{2} - \frac{1}{2}y - \frac{1}{2}\theta'(z) + \frac{1}{2} = x - y - \theta'(z) + 1$. Отождествим переменные x и y: $F_2(z) = F_1(x, x, z) = x - x - \theta'(z) + 1 = -\theta'(z) + 1 = 1 - \theta'(z)$.
Далее применим операцию суперпозиции для F и $F_2(z)$:

$F_3 = F(F_2(z), F_2(z), F_2(z)) = 1 - \theta'(z) - \frac{1}{2}(1 - \theta'(z)) - \frac{1}{2}\theta'(1 - \theta'(z)) + \frac{1}{2} = 1 - \theta'(z) - \frac{1}{2} + \frac{1}{2}\theta'(z) - \frac{1}{2}\theta'(1 - \theta'(z)) + \frac{1}{2} = 1 - \theta'(z) + \frac{1}{2}\theta'(z) - \frac{1}{2}\theta'(1 - \theta'(z))$.
Рассмотрим чему равно $\theta'(1 - \theta'(z))$. Так как $\theta'(1 - \theta'(z)) = \begin{cases} 0 & \text{при } z > 0 \\ 1 & \text{при } z \le 0 \end{cases}$.
и одновременно $1 - \theta'(z) = \begin{cases} 0 & \text{при } z > 0 \\ 1 & \text{при } z \le 0 \end{cases}$, значит $\theta'(1 - \theta'(z)) = 1 - \theta'(z)$. И следовательно $F_3 = 1 - \theta'(z) + \frac{1}{2}\theta'(z) - \frac{1}{2}(1 - \theta'(z)) = 1 - \theta'(z) + \frac{1}{2}\theta'(z) - \frac{1}{2} + \frac{1}{2}\theta'(z) = \frac{1}{2}$.

Теперь получим константу ноль. Для этого, подставим в функцию F_1 вместо переменной z, константу $\frac{1}{2}$: $F_4(x, y) = F_1(x, y, \frac{1}{2}) = x - y -$

$\theta^{'}(\frac{1}{2})+1 = x-y-1+1 = x-y$. Отожествляя переменные x и y, получим тождественный ноль. $F_4(x,x) = x - x = 0$.

Из $F_4(x,y)$ при помощи константы ноль получим функцию $-x$:

$F_5(y) = F_4(0,y) = 0 - y = -y$

$F_5(x) = -x$.

Затем из $F_4(x,y)$ и $F_5(y)$ суперпозицией выведем $x+y$:

$F_6(x,y) = F_4(x, F_5(y)) = x - (-y) = x + y$.

Теперь получим функцию $\frac{1}{2}x$. Для начала, получим константу 1. $F_6(\frac{1}{2},\frac{1}{2}) = \frac{1}{2}+\frac{1}{2} = 1$. Рассмотрим суперпозицию функций $F(x,y,z), F_5(x)$ и констант $0, 1$:

$$F(0, F_5(x), 1) = 0 - \frac{1}{2}(-x) - \frac{1}{2}\theta^{'}(1) + \frac{1}{2} = \frac{1}{2}x - \frac{1}{2} + \frac{1}{2} = \frac{1}{2}x.$$

И последнее, выведем $\theta(x)$.

$\theta(x) = 1 - \theta^{'}(-x)$.

Теорема доказана.

4.2. Построение базиса мощности k.

Теорема 3. *В классе кусочно-параллельных функций с двоично-рациональными коэффициентами BPP для любого наперед заданного $k \in \mathbb{N}$, существует базис из k элементов.*

. *Доказательство:* Приведем такие множества функций, составляющих базис BPP при $k < 5$.

Для $k = 4$ - базисом BPP является множество $B = \{\frac{1}{2}x, -x, x+y, \theta(x)\}$, по определению.

Докажем, что множество B действительно является базисом, в том смысле, что оно не избыточно. Т.е. нельзя выразить ни какую из входящих в него функций через остальные функции из B. Для этого поочередно будем исключать каждую функцию множества B и рассматривать, как будет меняться все множество функций составляющих класс BPP, используя общий вид функций принадлежащих классу BPP $(*)$. Если из B исключить $f_1 = \frac{1}{2}x$, то очевидно мы получим функции вида:
$(**) f(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_i + \dots + \theta(\omega_1) + \dots + \theta(\omega_s)$

$\theta(\omega_l) = \theta(x_1 + x_2 + \dots + x_{i'} + x^l_{i+1} + x^l_{i+2} + \dots + x^l_{j'-1} + \theta(\omega_{l1}) + \dots + \theta(\omega_{ls'}))$

$\theta(\omega_{li'}) = \theta(x_1 + x_2 + \dots + x_{i'} + x^{li'}_{i+1} + x^{li'}_{i+2} + \dots + x^{'}_{j'-1})$, т.е. множество линейных функций от n переменных с целыми коэффициентами и возможными разрывами 1го рода. Очевидно $\frac{1}{2}x$ не может быть представлена в виде $(**)$. Если из B исключить $f_2 = -x$, то мы получим функции вида:

$(***) f(x_1, x_2, \ldots, x_n) = \frac{k_1}{2^{m_1}}x_1 + \frac{k_2}{2^{m_2}}x_2 + \cdots + \frac{k_i}{2^{m_i}}x_i + \frac{k_{i+1}}{2^{m_{i+1}}}\theta(\omega_1) + \cdots + \frac{k_j}{2^{m_j}}\theta(\omega_s) + \frac{k_0}{2^{m_0}}$.

Где $k_0, k_1, \ldots, k_i, k_{i+1}, \ldots, k_j \in \mathbb{N}$,, т.к. коэффициент $\frac{k_i}{2^{m_i}}$ при x_i получается суперпозицией положительных функций $\frac{1}{2}x$ и $x + y$. И, следовательно, мы получаем класс линейных функций от n переменных, с положительными двоично-рациональными коэффициентами и возможными разрывами 1го рода. Функция $f_1 = -x$ не пренадлежит этому классу, т.к не может быть выражена через функции вида $(***)$.

Если из B исключить $f_3 = x + y$, то мы получим функции вида:

$f_1 = \frac{1}{2^m}x$,

$f_2 = \theta(\frac{1}{2^m}x)$,

$m_1, m_2, \ldots m_N \in \mathbb{N}, (N \leq n)$, т.е. функции зависящие от одной переменной.

Если из B исключить $f_4 = \theta(x)$, то очевидно мы получим функции вида:

$$f(x_1, x_2, \ldots, x_n) = \frac{k_1}{2^{m_1}}x_1 + \frac{k_2}{2^{m_2}}x_2 + \cdots + \frac{k_i}{2^{m_i}}x_i$$

Т.е. множество линейных непрерывных функций от n переменных, с двоично-рациональными коэффициентами. Получить разрывную функцию из непрерывных, используя операции суперпозиции и переименования, очевидно невозможно.

Следовательно множество B являющееся базисом для ВРР, действительно не избыточно при $|B| = 4$.

Для $k = 3$ - базисом ВРР является множество $\{-x, x + \frac{1}{2}y, \theta(x)\}$.

Для доказательства этого, выведем из данного множества функций, множество функций B. Операцией суперпозиции функции $F_1(x, y) = x + \frac{1}{2}y$ получим функцию $x+y$: $F_2(x, y) = F_1(F_1(x, y), y) = x + \frac{1}{2}y + \frac{1}{2}y = x + y$ Далее получим функцию $\frac{1}{2}x$:

$$F_2(F_1(x, y), -x) = x + \frac{1}{2}y - x = \frac{1}{2}y = \frac{1}{2}x.$$

В тоже время, ни одна функция не может быть удалена из данного, множества, что доказывается аналогично случаю $n = 4$.

Для $f_1 = x + \frac{1}{2}y$, это следует из того, что она необходима для образования функций: $f(x, y) = x + y$ и $f(x) = \frac{1}{2}x$.

Для $k = 2$ - базисом ВРР является множество $\{x - \frac{1}{2}y, \theta(x)\}$.

Докажем этот факт аналогично предыдущему. Сначала получим функцию $\frac{1}{2}x$ дважды применяя операцию суперпозиции к функции $F_1(x,y) = x - \frac{1}{2}$:

$F_2(x,y) = F_1(x, x - \frac{1}{2}y) = x - \frac{1}{2}(x - \frac{1}{2}y) = \frac{1}{2}(x + \frac{1}{2}y)$

$F_2(x - \frac{1}{2}y, y) = \frac{1}{2}(x - \frac{1}{2}y + \frac{1}{2}y) = \frac{1}{2}x.$

Затем получим функцию $x + y$:

$F_3(x,y) = F_1(x - \frac{1}{2}y, y) = x - \frac{1}{2}y - \frac{1}{2}y = x - y.$

$F_4(x,y) = F_1(x - y, y) = x - y - y = x - 2y.$

$F_5(x,y) = F_1(x, F_4(x,y)) = x - \frac{1}{2}(x - 2y) = x - \frac{1}{2}x + y = \frac{1}{2}x + y.$

$F_5(x, F_5(x,y)) = \frac{1}{2}x + \frac{1}{2}x + y = x + y.$

И в заключение суперпозицией функций $F_3(x,y), F_4(x,y)$ и $x + y$ и последующим отожествлением x и y получим функцию $-x$:

$F_3(F_4(x,y), x + y) = x - 2y - x + y = -y = -x$

Так как $F_1(x,y) = x - \frac{1}{2}y$ образует только непрерывные функции, $f(x) = \theta(x)$ необходима для образования ВРР, включающего в себя так же разрывные функции. В то же время функции $f(x) = \theta(x)$ не достаточно для образования базиса в ВРР, так как, например, с помощью нее не возможно получить двухместную функцию суммы $f(x) = x + y$, принадлежащую ВРР.

Для $k = 1$ - базисом ВРР является $\{x - \frac{1}{2}y - \frac{1}{2}\theta'(z) + \frac{1}{2}\}$, где $\theta'(z) = \begin{cases} 1 & \text{при } z > 0 \\ 0 & \text{при } z \leq 0 \end{cases}$, т.к. функция $F(x,y,z) = x - \frac{1}{2}y - \frac{1}{2}\theta'(z) + \frac{1}{2}$ является шефферовой в данном классе.

Теперь для $k \geq 5$ будем строить базис ВРР в виде:

$$\{\frac{1}{2}x, -x, p_1x + y, p_2x + y, \ldots, p_{k-3}x + y, \theta(x)\},$$

где НОД$(p_1, p_2, \ldots, p_{k-3}) = 1$, а НОД $(\bar{P}_i) \neq 1$,НОД$(\bar{P}_i) \neq 2m, \forall i, \bar{P}_i = (p_1, p_2, \ldots, p_{i-1}, p_{i+1}, \ldots, p_{k-3}), \forall m \in \mathbb{N}$.

Сначала докажем, что данная система функций порождает класс функций ВРР. Для этого нам достаточно получить функцию $x + y$. Рассмотрим числа $p_1, p_2, \ldots, p_{k-3}$. Так как НОД$(p_1, p_2, \ldots, p_{k-3}) = 1$, следовательно можно применив расширенный алгоритм Евклида найти такие числа $a_1, a_2, \ldots, a_{k-3} \in \mathbb{Z}$, что

$$a_1p_1 +_2 p_2 + \cdots + a_{k-3}p_{k-3} = 1$$

Произведя $|a_1|$ раз операцию суперпозиции функции $F_1(x,y) = p_1x + y$ вида $F_1(x, F_1(x,y))$, получим функцию $F_1^2(x,y) = a_1p_1x + y$, затем произведем a_2 раз операцию суперпозиции функции $F_2(x,y) = p_2x + y$ вида

$F_1(x, F_1(x,y))$, получим функцию $F_2^2(x,y) = a_2p_2x + y$ и так получим $k-3$ функции, вида $F_i^2(x,y) = a_ip_ix + y$, где $i = \overline{1, k-3}$. Далее произведем суперпозицию $F_+^1(x,y) = F_2^1(x, F_2^2(x,y)) = a_1p_1x + a_2p_2x + y$. И так последовательно будем проводить суперпозиции вида:

$F_+^i(x,y) = F_+^{i-1}(x, F_1^{i+1}(x,y))$, для всех $k-3$ функций $F_i^2(x,y)$. В итоге получим функцию $F_+(x,y) = F_+^{k-4}(x,y) = a_1p_1x + a_2p_2x + \cdots + a_{k-3}p_{k-3}x + y = (a_1p_1 + a_2p_2 + \cdots + a_{k-3}p_{k-3})x + y = x + y$.

Докажем не избыточность этого множества. Необходимость функций $\frac{1}{2}x, -x, \theta(x)$ доказывается аналогично случаю $k = 4$. Все функции $p_1x + y, p_2x + y, \ldots, p_{k-3}x + y$ являются необходимыми, в силу того, что числа $p_1, p_2, \ldots, p_{k-3}$ попарно не взаимно-просты, т.е., НОД$(\bar{P}_i) \neq 1$,а также - НОД$(\bar{P}_i) \neq 2m$. Рассмотрим множество функций $B' = B \backslash p_ix + y$. Пусть НОД$(\bar{P}_i)) = c$, тогда из алгоритма Евклида, будет следовать, что при суперпозиции функций принадлежащих множеству B мы получим сумму вида: $f_+(x,y) = cx + y$, где $c \neq 1$. Причем, так как НОД$(\bar{P}_i) \neq 2m$, т.е. $c \neq 2m$, единицу не возможно получить применяя m раз суперпозицию $f_+(\frac{1}{2}x, y)$.
Теорема доказана.

Например, для $k = 6$, базис ВРР составляет множество функций $\frac{1}{2}x, -x, 231x + y, 165x + y, 105x + y, 385x + y, \theta(x)$.

Автор выражает искреннюю признательность Часовских А.А. за постановку задачи, а также за обсуждение результатов работы за ценные советы и замечания.

Список литературы

[1] Половников В. С. Об оптимизации структурной реализации нейронных сетей. Диссертация на соискание ученой степени кандидата физико-математических наук — Москва, 2006.

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

[3] Хайкин. С. Нейронные сети: полный курс // 2-е издание. Вильямс, 2006. // Вестн. Моск. ун-та. Матем. Механ. — 2016. — № 4. — С. 12–17.

[4] Кан А. Н. Вопросы выразимости в классе нейронных функций //Интеллектуальные системы — том 19, — Выпуск 1. — 2015.

[5] Яблонский С. В. Введение в дискретную математику. . — М.: Изд-во Наука, 1986.

[6] Кудрявцев В. Б. Функциональные системы. . — М.: Изд-во МГУ, 1982.

On a class of neural functions with binary-rational parameters

Agafonova M. V.

The paper deals with a class of neural piecewise-parallel functions with binary-rational coefficients. It is proved that for any function in the class of piecewise-parallel functions there exists a function in the class of piecewise-parallel functions with binary-rational coefficients that approximates it with a predetermined accuracy. Also, for the class of functions under consideration, it is shown that there exist bases consisting of arbitrary number of elements, in particular, a Scheffer function is found.

Keywords: class of piecewise-parallel functions, class of neural functions with binary-rational coefficients, superposition operations, Scheffer function, bases.