УДК 681.14

Ф.Е. Гече, В.М. Коцовський, С.А. Ковальов
Ужгородський нац. ун-т

ПРО ЗБІЖНІСТЬ СПЕКТРАЛЬНОГО АЛГОРИТМУ НАВЧАННЯ НЕЙРОННИХ ЕЛЕМЕНТІВ.

В роботі розглядаються узагальненні нейронні елементи відносно системи характерів, наводиться алгоритм навчання цих елементів та доводиться його збіжність.

Generalized neural devices over the character set have been studied in the paper. Learning algorithm has been given and its convergence has been proved in the paper.

Нейронні елементи (НЕ) використовуються при побудові нейромереж, які є ефективним механізмом для їх зв'язування широкого кола задач класифікації об'єктів, розпізнавання образів, стиску інформації, прогнозування поведінки динамічних систем, наближення і екстраполяції функцій [1-4]. У роботі вивчаються НЕ відносно системи характерів, розгляд яких дозволяє значно розширити клас бульових нейрофункцій, описується та обґрунтовується метод навчання таких нейронних елементів.

Нехай $H_2 = \{-1,1\}$, $G_n = H_2 \times \ldots \times H_2$ – n-а декартова степінь множини H_2. Дискретні функції вигляду $f: G_n \to H_2$ будемо називати бульовими функціями в алфавіті $\{-1,1\}$. Визначимо відображення $\chi_j : G_n \to H_2$ наступним чином $\chi_j(a) = a_1^{j_1} \cdot a_2^{j_2} \cdot \ldots \cdot a_n^{j_n}$, де $a = (a_1,\ldots,a_n) \in G_n$, $j = j_1 2^{n-1} + j_2 2^{n-2} + \ldots + j_n$ $\left(j_i \in \{0,1\},\ i = \overline{1,n}\right)$. Відображення χ_j, $j = 0,1,\ldots,2^n - 1$ будемо називати характерами групи G_n над полем R [4]. На множині $R\setminus\{0\}$ визначимо знакову функцію Rsign так:

$$\operatorname{Rsign} x = \begin{cases} 1, & \text{якщо } x > 0, \\ -1, & \text{якщо } x < 0. \end{cases}$$

Нехай $X = \{\chi_{i_1},\ldots,\chi_{i_m}\}$ – деяка система характерів, $w = (w_1,\ldots,w_m) \in R^m$. Якщо для бульової функції $f: G_n \to H_2$ виконується умова

$$\text{для всіх } a = (a_1,\ldots,a_n) \in G_n \ f(a) = \operatorname{Rsign} \sum_{j=1}^{n} w_j \chi_{i_j}(a), \tag{1}$$

то будемо казати, що відносно системи характерів X бульова функція f реалізується на НЕ з ваговим вектором w. Якщо для бульової функції f можна вказати принаймні один такий ваговий вектор w, для якого виконується умова (1), то функцію f будемо називати *X-нейрофункцією*. При цьому для вагового вектора w має виконуватися умова для всіх $a \in G_n$ $(w, \chi(a)) \neq 0$, де $(w, \chi(a)) = w_1\chi_{i_1}(a) + \ldots + w_m\chi_{i_m}(a)$ – скалярний добуток векторів w і $\chi(a) = (\chi_1(a),\ldots,\chi_m(a))$. Такі вагові вектори будемо називати *X-допустимими*. Позначимо через $W_X(f)$ множину вагових векторів всіх НЕ, на яких реалізується *X-нейрофункція* f.

Під процесом навчання НЕ для бульової *X-нейрофункції* f будемо розуміти процес побудови скінченої послідовності векторів

$$w^0, w^1, \ldots, w^l, \tag{2}$$

такої, що функція f реалізується на НЕ з ваговим вектором w^l.

Для бульової функції $f: G_n \to H_2$ відносно системи характерів X визначимо характеристичний вектор $s(f) = (s_1,\ldots s_m)$ так: $s_j = \sum_{a \in G_n} \chi_{i_j}(a) f(a)$, $j = \overline{1,m}$. Будемо вважати, що перші $k+1$ членів $w^0, w^1, \ldots, w^k$ послідовності (2) вже відомі і нехай w^k – m-вимірний *X-допустимий* дійсний вектор. Якщо функція f реалізується на НЕ з ваговим вектором w^k, то процес навчання вважаємо завершеним. Припустимо, що f не реалізується на НЕ з ваговим вектором w^k. Опишемо алгоритм, за яким можна побудувати такий вектор w^{k+1}, для якого справджується нерівність

$$\left\| w^{k+1} - w \right\| < \left\| w^k - w \right\|,$$

де $\|x\| = \sqrt{(x,x)}$ – звичайна евклідова норма в просторі R^m). Нехай $z^k = w^{k+1} - w^k$ – вектор приросту, $f^k(x_1,\ldots x_n)$ – функція, яка відносно системи характерів X реалізується на НЕ з ваговим вектором w^k. Вектор приросту z^k будемо шукати у вигляді

$$z^k = t_k \alpha_k^0 \left(s_X(f) - s_X(f^k)\right), \tag{3}$$

де

$$\alpha_k^0 = \frac{\left(w^k, s_X(f^k) - s_X(f)\right)}{\left\|s_X(f) - s_X(f^k)\right\|^2},$$

$$0 < t_k < 2\left(1 + \frac{\left(w, s_X(f) - s_X(f^k)\right)}{\left(w^k, s_X(f^k) - s_X(f)\right)}\right). \tag{4}$$

Число t_k назвемо нормуючим множником приросту.

Зауваження 1. При виборі вектора приросту z^k у вигляді (3) необхідно вимагати, щоб для всіх $a \in G_n$ $\left(w^{k+1}, \chi(a)\right) \neq 0$, бо інакше вектор w^{k+1} не є *X-допустимим* і на НЕ з таким ваговим вектором не можна реалізувати жодну цілком визначену бульову функцію. Цього можна завжди досягти за рахунок незначних змін вектора приросту, змінивши відповідним чином нормуючий множник приросту t_k так, щоб він не вийшов за межі допустимих значень (4).

Можна показати, що якщо $t_k \geq 1$, то виконується нерівність $s_X(f^{k-1}) \neq s_X(f^k)$.

Зауваження 2. Попередня властивість не виключає того, що знайдуться такі натуральні числа k і l $(k < l)$, для яких має місце рівність $s_X(f^k) = s_X(f^l)$. В цьому випадку можна провести дослідження граничного циклу по тій самій схемі, яка була застосована у [5] для порогових функцій.

Розглянемо два підходи до вибору нормуючих множників t_k. Перший підхід запропонований у роботі [5] і ґрунтується на наступній теоремі.

Теорема 1. *Нехай f – бульова функція, w^0 – X-допустимий ваговий вектор, прирости для послідовності $\{w^k\}$ визначаються згідно (3), нормуючий множник t_k вибирається з проміжку $[t_0, 2]$, де $0<t_0<2$, $f^k \neq f$, $k=0,1,\ldots$. Тоді якщо знайдеться таке $\varepsilon > 0$, що для всіх $k \in N$ і для всіх $a \in G_n$ виконується умова*

$$\left|\left(w^k, \chi(a)\right)\right| \geq \varepsilon, \tag{5}$$

то функція f не є X-нейрофункцією.

Зауваження 3. Якщо бульова функція f є *X-нейрофункцєю* і для послідовності $\{w^k\}$ умова (5) не виконується, то не можна стверджувати, що $w^* \in W_X(f)$. В цьому випадку може виявитися, що вектор w^* уже не є X-допустимим і належить замиканню множини $W_X(f)$.

На основі теореми 1 можна запропонувати таку модифікацію алгоритму навчання при умові, що нормуючий множник t_k вибирається з проміжку $[t_0, 2]$. Вибираємо деяке достатньо мале число $\varepsilon > 0$ і здійснюємо процес навчання до тих пір, поки не почне виконуватися одна з наступних умов:

1. $s_X(f^k) = s_X(f)$. Тоді робимо висновок, що бульова функція f відносно системи характерів X реалізується на НЕ з ваговим вектором w^k і процес навчання завершено.
2. В області характеристичних векторів утворився граничний цикл. Тоді діємо згідно зауваження 2.
3. Порушилася нерівність (5). Тоді припиняємо процес навчання або вносимо в нього якісь зміни (змінюємо початкове наближення, нормуючі множники t_k на кількох останніх кроках алгоритму або зменшуємо ε).

Тестування спектрального алгоритму навчання показує, що випадок 3 при навчанні *X-нейрофункцій* майже не зустрічається. Проте можна запропонувати правило вибору нормуючих множників, при застосуванні якого до навчання *X-нейрофункцій* спектральний алгоритм навчання через скінчену кількість кроків завжди приводить до вагового вектора w' НЕ, який реалізує цю функцію. Для цього будемо вибирати множники t_k таким чином:

$$t_k = 2 + \frac{1}{\left(w^k, s_X(f^k) - s_X(f)\right)}. \tag{6}$$

Теорема 2. *Якщо бульова функція f є X-нейрофункцією, послідовність X-допустимих вагових векторів $\{w^k\}$ будується за допомогою приростів виду (9), нормуючі множники t_k вибираються згідно (6) з урахуванням зауваження 1, то процес навчання завершується через скінчену кількість кроків на деякому векторі $w' \in W_X(f)$.*

Зауваження 4. Можливий варіант алгоритму, в якому поєднуються обидва підходи до вибору нормуючих множників: вибираємо деяке достатньо мале число $\varepsilon > 0$, здійснюємо алгоритм навчання, поклавши $t_k=2$ (з урахуванням зауваження 2) до тих пір, поки виконується умова (5), а в разі її порушення продовжуємо процес навчання, вибираючи t_k за формулою (6).

Зауваження 5. Можна запропонувати інші модифікації алгоритму навчання узагальнених нейронних елементів над системою характерів, які схожі до алгоритмів навчання персептронів, наведених у роботах [3] і [6]. У цих алгоритмах не вимагається обчислення спектральних параметрів бульових функцій. Крім того, вони можуть бути застосовані для навчання неповністю заданих бульових функцій. Для цих алгоритмів можна провести оцінку максимального числа ітерацій, потрібних для навчання. Тому вони можуть використовуватися для перевірки того, чи є задана функція *fX-нейрофункцією*.

Література

1. *Уоссермен Ф. Нейрокомпьютерная техника: теория и практика.– М.:Мир,1992.*
2. *Грицик В.В., Гече Ф.Е. Реалізація бульових та багатозначних бульових функцій на нейронних елементах // Доповіді НАН України №5, 2004, С. 65-68.*
3. *Айзенберг Н.Н., Иваськив Ю.Л. Многозначная пороговая логика. – Київ: Наукова думка, 1977.*
4. *Ван дер Варден Б. Л. Алгебра. – М.: Наука, 1979.*
5. *Дертоузос М. Пороговая логика. – М.: Мир, 1967.*
6. *Розенблатт Ф. Принципы нейродинамики. – М. Мир, 1965.*