

А.С. Овечкин

Применение неклассических моделей нейронов в задачах ассоциативной памяти на КНС-структуре

В статье рассматривается использование неклассических нейроэлементов в клеточной нейроструктуре. Приводятся алгоритмы обучения КНС на этих элементах для реализации задачи ассоциативной памяти.

Одними из наиболее важных в практическом применении нейросетевых структур являются *клеточные нейронные сети (КНС)*. Как известно, отличительной чертой данной конфигурации является локальность межэлементных связей. Фактически в плоском случае *КНС* представляет собой прямоугольную решетку, в узлах которой находятся нейроны подобные элементы, каждый из которых имеет синаптическую связь с соседями из рецептивной окрестности. Дальнейшее рассмотрение будет производиться для *окрестности Мура*, представляющей собой 8 ближайших соседей элемента, состояния элементов в узлах будут предполагаться $\in(-1; +1)$.

В классическом варианте узлы решетки представляют собой нейроны МакКаллока-Питтса, т.е. простейшие сумматоры (без учета внешнего сигнала) с нелинейной функцией активации:

$$H(x) = \sum x_i a_i; \quad (1)$$

$$f(H(x)) = \begin{cases} 1, & \text{если } H(x) - J < 0, \\ -1, & \text{если } H(x) - J \leq 0, \end{cases} \quad (2)$$

где $x = (x_1, \dots, x_n)$ - вектор входов нейрона, $a = (a_1, \dots, a_n)$ - вектор синаптических весов, $H(x)$ - уровень активации, J - порог активации нейрона, $f(H(x))$ - функция активации. Данная модель нейрона, как известно, обладает ограниченными классифицирующими свойствами, вследствие чего представляется разумным ее усложнить, сохраняя локальность связей *КНС*.

Сравнительно недавно были открыты пары согласованно функционирующих биологических нейронов, один из которых является тормозящим, а

другой - возбуждающим. Нейроны функционируют в альтернативном режиме: в возбужденное состояние (генерация сигнала) переходит либо первый, либо второй нейрон. Абстрагируясь от физиологических деталей, такую пару нейронов можно рассматривать как один нейронный элемент, обладающий специфическими свойствами. Такой нейрон получил название *альтернативного* [1] (*A-нейрон*).

Пусть $x = (x_1, \dots, x_n)$ - вектор входов, $a = (a^+, a^-)$ - вектор синаптических весов *A-нейрона*, причем $a^+ = (a_1^+, \dots, a_n^+)$; $a^- = (a_1^-, \dots, a_n^-)$, тогда уровень активации *A-нейрона* без учета внешнего сигнала выглядит следующим образом:

$$H(x) = \sum x_i (a_i^+ \theta(x_i) + a_i^- \theta(-x_i)),$$

$$\theta(z) = \begin{cases} 1, & \text{если } z \geq 0, \\ 0, & \text{если } z < 0. \end{cases} \quad (4)$$

Еще одним примером усложнения классического нейрона может служить *нейрон с корреляционными синапсами*, или *нейрон высшего порядка*. Принципиально устройство данной модели можно пояснить на простейшем примере. Предположим, что некоторый нейрон имеет два входа - x_1 и x_2 с соответствующими синаптическими весами a_1 и a_2 . Тогда, рассматривая и возможное взаимодействие между входами x_1 и x_2 , на нейрон оказывается дополнительное воздействие, вычисляемое как $a_{12}x_1x_2$. Величина a_{12} называется корреляционным входом нейрона, а a_{12} - корреляционным синапсом, соответствующим этому входу. Будем говорить, что такой корреляционный вход (и соответствующий ему корреляционный синапс) имеет 2-й порядок (или степень), т.к. рассматривается взаимодействие двух входов. Таким образом, можно сказать, что классический нейрон МакКаллока-Питтса обладает лишь входами первого порядка, т.е. является частным случаем нейрона с корреляционными синапсами. (Можно также считать нулевыми синапсы порядков выше первого у нейрона МакКаллока-Питтса.) Итак, в общем случае уровень активации нейрона с корреляционными синапсами, или нейрона высшего порядка, без учета внешнего воздействия вычисляется так:

$$H(x) = \sum_{i_1} a_{i_1}^{(1)} x_{i_1} + \sum_{i_1 < i_2} a_{i_1, i_2}^{(2)} x_{i_1} x_{i_2} + \dots + \sum_{i_1 < \dots < i_N} a_{i_1, \dots, i_N}^{(N)} x_{i_1} \dots x_{i_N}, \quad (5)$$

где N - количество входов первого порядка.

Рассмотрим далее реализацию ассоциативной памяти (АП) на рассмотренных моделях нейронов. Как известно, постановка задачи АП в общем случае имеет вид:

Пусть задан набор из t эталонов (n -мерных векторов). Требуется построить (т.е. сконструировать и обучить) сеть, которая при предъявлении на вход произвольного образа-вектора давала бы на выходе наиболее похожий эталон.

Таким образом, выбирая для реализации АП клеточную структуру, главной проблемой становится проблема обучения сети. Дело в том, что широко известный и простой в реализации алгоритм Хэбба, вообще говоря, не обеспечивает важнейшее динамическое свойство обученной АП - индивидуальную устойчивость эталонов. Ввиду этого описанная в работе [2] итеративная процедура обучения персептрона МакКаллока-Питтса была модифицирована для обучения клеточных АП на нейронах высших порядков и А-нейронах.

В классическом варианте данная процедура обучения выглядит следующим образом. Для нейрона α введем величины: вектор приведенного соседства и вектор синаптических весов от соседей

$$D_\alpha = x_\alpha(x_{i_1}, \dots, x_{i_q}); A_\alpha = (a_{i_1, \alpha}, \dots, a_{i_q, \alpha}),$$

где q - количество соседей нейрона α ; $i_1, \dots, i_q \in N(\alpha)$, т.е. соседи нейрона α .

Итак, пусть имеется набор из M эталонов - векторов из R^n т.е. набор $(x^{(1)}, \dots, x^{(M)})$, где $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$, а его компоненты $x_i^{(k)} \in (-1; +1)$.

Шаг 1. Из данных эталонов организуем бесконечную последовательность $x^{(1)}, \dots, x^{(M)}, x^{(1)}, \dots, x^{(M)}, \dots$, полученную из исходного множества эталонов его циклическим повторением. Период обучения, в течение которого на вход сети были поданы $x^{(1)}, \dots, x^{(M)}$, будем называть макроитерацией обучения.

Шаг 2. Начальные (в момент времени $t = 0$) значения компонент векторов A_α^0 выбираем произвольно для всех нейронов α нашей сети. Параметр обучения $\beta_\alpha^0 = 0$, $\tau = 0$ - счетчик макроитераций.

Шаг 3. Начало макроитерации, $\mu_\alpha^\tau(t) = \infty$; выполнение *шага 4* для всего исходного множества эталонов.

Шаг 4. Весовые векторы изменяются в соответствии со следующей итеративной процедурой:

$$A_\alpha^{t+1} = \begin{cases} A_\alpha^t, & \text{если } (D_\alpha^{(t)}, A_\alpha^t) > \beta_\alpha^\tau, \\ A_\alpha^t + D_\alpha^{(t)}, & \text{если } (D_\alpha^{(t)}, A_\alpha^t) \leq \beta_\alpha^\tau. \end{cases} \quad (6)$$

$$\mu_\alpha^\tau(t+1) = \min(\mu_\alpha^\tau(t), (D_\alpha^{(t)}, A_\alpha^t)). \quad (7)$$

Шаг 5. Если в результате макроитерации появляется условие останова, то выполняется останова; в противном случае вычисляется новый параметр

обучения $\beta_{\alpha}^{\tau+1} = \max(\beta_{\alpha}^{\tau}, \mu_{\alpha}^{\tau}(t))$. В качестве условия останова можно взять: 1) ограничение числа макроитераций; 2) ограничение нормы весового вектора A_{α} .

Теперь применим эту процедуру для обучения КНАП на нейронах высших порядков. Действительно, это можно сделать без каких-либо изменений или модификаций приведенного выше классического алгоритма. В самом деле, поскольку совместное состояние группы нейронов (численно равное произведению состояний нейронов, входящих в группу), являющихся соседями некоторой клетки, можно рассматривать как состояние некоего виртуального соседа, то фактически алгоритм обучения для данной модели отличается от описанного выше только значительным увеличением числа соседей, что никак не сказывается на его формальном описании.

В следствие этого мы не станем останавливаться на алгоритме для данной модели, а рассмотрим более подробно итеративный алгоритм, адаптированный для обучения КНАП на А-нейронах.

Итак, для нейрона α введем величины: вектор приведенного соседства и вектор синаптических весов от соседей

$$\begin{aligned} D_{\alpha}^{+} &= x_{\alpha}^{+}(x_{i_1}^{+}, \dots, x_{i_q}^{+}); D_{\alpha}^{-} = x_{\alpha}^{-}(x_{i_1}^{-}, \dots, x_{i_q}^{-}), \\ A_{\alpha}^{+} &= (a_{i_1, \alpha}^{+}, \dots, a_{i_q, \alpha}^{+}); A_{\alpha}^{-} = (a_{i_1, \alpha}^{-}, \dots, a_{i_q, \alpha}^{-}), \end{aligned}$$

где q - количество соседей нейрона α ; $i_1, \dots, i_q \in N(\alpha)$, т.е. соседи нейрона α .
Причем

$$x_{i_k}^{+} = \begin{cases} x_{i_k}^{+}, & \text{если } x_{i_k}^{+} > 0, \\ 0, & \text{если } x_{i_k}^{+} \leq 0, \end{cases} \quad (8)$$

$$x_{i_k}^{-} = \begin{cases} 0, & \text{если } x_{i_k}^{-} > 0, \\ x_{i_k}^{-}, & \text{если } x_{i_k}^{-} \leq 0. \end{cases} \quad (9)$$

Итак, пусть имеется набор из M эталонов - векторов из R^n т.е. набор $(x^{(1)}, \dots, x^{(M)})$, где $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$, а его компоненты $x_i^{(k)} \in (-1; +1)$.

Шаг 1. Из данных эталонов организуем бесконечную последовательность $x^{(1)}, \dots, x^{(M)}, x^{(1)}, \dots, x^{(M)}, \dots$, полученную из исходного множества эталонов его циклическим повторением. Период обучения, в течение которого на вход сети были поданы $x^{(1)}, \dots, x^{(M)}$, будем называть макроитерацией обучения.

Шаг 2. Начальные (в момент времени $t = 0$) значения компонент векторов $A_{\alpha}^{+,0}$ и $A_{\alpha}^{-,0}$ выбираем произвольно для всех нейронов α нашей сети. Параметры обучения $\beta_{\alpha}^{+,0} = 0$ и $\beta_{\alpha}^{-,0} = 0$, $\tau = 0$ - счетчик макроитераций.

Шаг 3. Начало макроитерации. $\mu_{\alpha}^{+,\tau}(t) = \infty$, $\mu_{\alpha}^{-,\tau}(t) = \infty$; выполнение шага 4 для всего исходного множества эталонов.

Шаг 4. Весовые векторы изменяются в соответствии со следующей итеративной процедурой:

$$A_{\alpha}^{+,t+1} = \begin{cases} A_{\alpha}^{+,t}, & \text{если } (D_{\alpha}^{+,t}, A_{\alpha}^{+,t}) > \beta_{\alpha}^{+,\tau}, \\ A_{\alpha}^{+,t} + D_{\alpha}^{+,t}, & \text{если } (D_{\alpha}^{+,t}, A_{\alpha}^{+,t}) \leq \beta_{\alpha}^{+,\tau}. \end{cases} \quad (10)$$

$$\mu_{\alpha}^{+,\tau}(t+1) = \min(\mu_{\alpha}^{+,\tau}(t), (D_{\alpha}^{+,t}, A_{\alpha}^{+,t})) \quad (11)$$

$$A_{\alpha}^{-,t+1} = \begin{cases} A_{\alpha}^{-,t}, & \text{если } (D_{\alpha}^{-,t}, A_{\alpha}^{-,t}) > \beta_{\alpha}^{-,\tau}, \\ A_{\alpha}^{-,t} + D_{\alpha}^{-,t}, & \text{если } (D_{\alpha}^{-,t}, A_{\alpha}^{-,t}) \leq \beta_{\alpha}^{-,\tau}. \end{cases} \quad (12)$$

$$\mu_{\alpha}^{-,\tau}(t+1) = \min(\mu_{\alpha}^{-,\tau}(t), (D_{\alpha}^{-,t}, A_{\alpha}^{-,t})) \quad (13)$$

Шаг 5. Если в результате макроитерации появляется условие останова, то выполняется останов; в противном случае вычисляются новые параметры обучения $\beta_{\alpha}^{+,\tau+1} = \max(\beta_{\alpha}^{+,\tau}, \mu_{\alpha}^{+,\tau}(t))$ и $\beta_{\alpha}^{-,\tau+1} = \max(\beta_{\alpha}^{-,\tau}, \mu_{\alpha}^{-,\tau}(t))$.

Итак, для экспериментального изучения и сравнения описанных выше моделей нейронов были созданы программные модели КНС на базе нейронов МакКаллока-Питтса, нейронов высших порядков и А-нейронов размером 30x30 элементов с окрестностью Мура радиусов 1 и 2, которые использовались для реализации АП. Каждая сеть была обучена на наборе из 5, 10 и 15 черно-белых образов. В качестве образов использовались буквы латинского алфавита, выбранные произвольно. В процессе распознавания на вход сети подавались образы из обучающей последовательности, искаженные случайным шумом. Искажения заключались в зачернении с вероятностью пикселей входного паттерна. По истечении определенного времени в КНС устанавливался устойчивый режим работы, в котором измерялось расстояние Хемминга от результата до исходного неискаженного образа, т.е. нормированное коэффициентом $N=30 \times 30$ число различий между эталонным и полученным образом, а также среднеквадратичное отклонение получаемых образов. В результате проведенных экспериментов можно с уверенностью утверждать, что работа АП на А-КНС и корреляционной КНС превосходит АП на классической КНС по всем параметрам:

- 1) имеет меньший среднеквадратичный разброс, что говорит о ее устойчивости,
- 2) имеет значительно меньшие средние показатели расстояний Хемминга, причем при использовании радиуса окрестности 2 все показатели демонстрируют значительное превосходство неклассических моделей нейронов.

Литература

1. *Панкратов В.А.* Некоторые модели локальных нейросетей и их исследование. Ярославль, 2000 // Дис. канд. физ.-мат. наук.
2. *Пудов С.Г.* Обучение клеточно-нейронной ассоциативной памяти // Автометрия. Новосибирск, 1997. №2. 1997. С. 107-120.

Ярославский государственный университет им. П.Г. Демидова