

АЛГОРИТМ ОБУЧЕНИЯ РЕКУРРЕНТНОГО ПЕРСЕПТРОНА С СИММЕТРИЧНЫМИ СВЯЗЯМИ

Борило Илья Анатольевич

аспирант Томского Государственного университета, РФ, г. Томск

E-mail: boriloilya@yandex.ru

Слядников Евгений Евгеньевич

д-р физ.-мат. наук, профессор Томского Государственного университета, РФ, г. Томск

E-mail: opi@hq.tsc.ru

LEARNING ALGORITHM OF RECURRENT PERCEPTRON WITH SYMMETRIC CONNECTION

Borilo Ilya Anatolyevich

phD student of Tomsk State University, Russia Tomsk

Slyadnikov Evgeniy Evgenievich

dr. phys.-Math. Science, professor of Tomsk State University, Russia Tomsk

АННОТАЦИЯ

В статье рассмотрен рекуррентный персептрон с симметричными связями. Приведены несколько модификаций алгоритма обратного распространения ошибки, позволяющие проводить обучение предложенной модели рекуррентного персептрона. На примере предсказания элементов процесса Маккея-Гласса проведено экспериментальное сравнение предложенных модификаций.

ABSTRACT

Recurrent perceptron with symmetric connection is described in the article. Modification back propagation algorithm, which could be used for learning of described recurrent perceptron, is described. Learning efficiency of Mackey-Glass process prediction is compared.

Ключевые слова: рекуррентный персептрон; алгоритм обратного распространения ошибки; предсказание последовательности.

Keywords: recurrent perceptron; algorithms of backward propagation of errors; series prediction.

В тексте статьи приведено описание рекуррентного персептрона с симметричными связями. Предложенная модель интересно тем, что она может быть использована для моделирования обработки информации в микротрубочке цитоскелета [1—4]. При этом возникает задача обучения предложенного персептрона для аппроксимации заданной зависимости. Новизна поставленной задачи состоит в том, что на коэффициенты связи нейронов накладывается дополнительное ограничение: любые два нейрона связаны прямой и обратной связью, причем значения коэффициентов связи одинаковы — при этом данный персептрон не является аттракторным, т. е. рассматриваемый персептрон не эквивалентен персептрону Хопфилда. Разработка данного алгоритма является целью данной работы. В работе рассматривалась только аппроксимация числовых последовательностей (7), значение каждого элемента которых зависит от конечно числа предыдущих элементов.

$$s(t) = f(s(t-1), s(t-2), \dots, s(t-N)) \quad (1)$$

где: $s(t)$ — элемент последовательности с индексом t ;

$f(s(t-1), s(t-2), \dots, s(t-N))$ — функция, определяющая следующий элемент последовательности по N предыдущим элементам данной последовательности.

Далее будем рассматривать аппроксимацию функции (1) с помощью рекуррентного персептрона с симметричными связями. В данной работе будем рассматривать простую топологию такого персептрона. Персептрон будет состоять из двух слоев, причем в выходном слое будет только один нейрон. Значения нейронов в скрытом слое будет определяться выражениями (2) и (3); значения нейронов в выходном слое будут определяться выражениями (4) и (5).

$$v_i(t) = \sum_{j=0}^N w_{ij} \cdot x_j(t) + w_i \cdot y(t-1) \quad (2)$$

$$y_i(t) = th(v_i(t)) \quad (3)$$

$$v(t) = \sum_{i=1}^m w_i \cdot y_i(t) \quad (4)$$

$$y(t) = th(v(t)) \quad (5)$$

где: $v_i(t)$ — потенциал активации нейрона в скрытом слое с индексом i ;

$v(t)$ — потенциал активации нейрона в выходном слое ;

$x_j(t)$ — компонента входного вектора с индексом j , где $x_j(t) = d(t - j)$,

$x_0(t) = 1$;

$y_i(t)$ — значение нейрона в скрытом слое с индексом i ;

$y(t)$ — значение нейрона в выходном слое, которое будет оценкой элемента последовательности $d(t)$ с индексом t , вычисленной персептроном;

w_{ij} — коэффициент связи компоненты входного вектора с индексом j и нейрона в скрытом слое с индексом i ;

w_i — коэффициент связи нейрона в скрытом слое с индексом i с нейроном выходным слое, одновременно это коэффициент связи предыдущего значения выходного нейрона с нейроном в скрытом слое с индексом i ;

m — число нейронов в скрытом слое.

Процесс обучения такого персептрона, чтобы он аппроксимировал заданную последовательность описываемую (1), будет состоять в последовательной коррекции значений коэффициентов связи персептрона методом градиентного спуска, где изменение значение коэффициента связи будет определяться выражением (6).

$$\Delta w = -\eta \cdot \frac{d(s(t) - y(t))^2 / 2}{dw} \quad (6)$$

где: η — параметр скорости обучения.

Оценка значения (6) вычисляется с помощью алгоритма обратного распространения ошибки [5]. Значение (6) для коэффициента w_{ij} будет равно (7); значение (6) для коэффициента w_i будет равно (8).

$$\Delta w_{ij} = \eta \cdot (s(t) - y(t)) \cdot \frac{dy(t)}{dv(t)} \cdot w_i \cdot \frac{dy_i(t)}{dv_i(t)} \cdot x_j \quad (7)$$

$$\Delta w_i = \eta \cdot (s(t) - y(t)) \cdot \frac{dy(t)}{dv(t)} \cdot \left(y_i(t) + w_i \cdot \frac{dy_i(t)}{dv_i(t)} \cdot y(t-l) \right) \quad (8)$$

Также оценку значений (6) для коэффициентов w_{ij} и w_i можно вычислить, используя (9) и (10) соответственно.

$$\Delta w_{ij} = \eta \cdot (s(t) - y(t)) \cdot \frac{dy(t)}{dv(t)} \cdot \left(w_i \cdot \frac{dy_i(t)}{dv_i(t)} \cdot x - y_i(t) \cdot \delta_{jl} \right)_j \quad (9)$$

$$\Delta w_i = \eta \cdot (s(t) - y(t)) \cdot \frac{dy(t)}{dv(t)} \cdot y_i(t) \quad (10)$$

где: δ_{jl} — дельта функция.

Т. к. $y(t-l) = s(t-l) + \varepsilon$ в (2), где ε — погрешность аппроксимации, то чем больше будет значение w_i , тем больше будет погрешность оценки $y(t)$. Поэтому логично наложить ограничение на значения коэффициентов w_i , $w_i \in [-w_c, w_c]$, где w_c — некоторая константа.

Далее будем рассматривать четыре возможные реализации алгоритма обучения рекуррентного персептрона: с использованием (7) и (8) (BP7_8), с использованием (7) и (8) с ограничением значений коэффициентов (BP7_8_bord), с использованием (9) и (10) (BP9_10), с использованием (9) и (10) с ограничением значений коэффициентов (BP9_10_bord).

Экспериментальное сравнение предложенных реализаций было приведено на задаче предсказания последовательности значений процесса Маккея-Гласса [6]. Данная последовательность описывается выражением:

$$\frac{ds(t)}{dt} = \frac{a \cdot s(t - \tau)}{1 + s^{10}(t - \tau)} - b \cdot s(t) \quad (10)$$

где $a=0.1$, $b=0.2$, $\tau=30$. Пример такой последовательности приведен на рисунке 1.

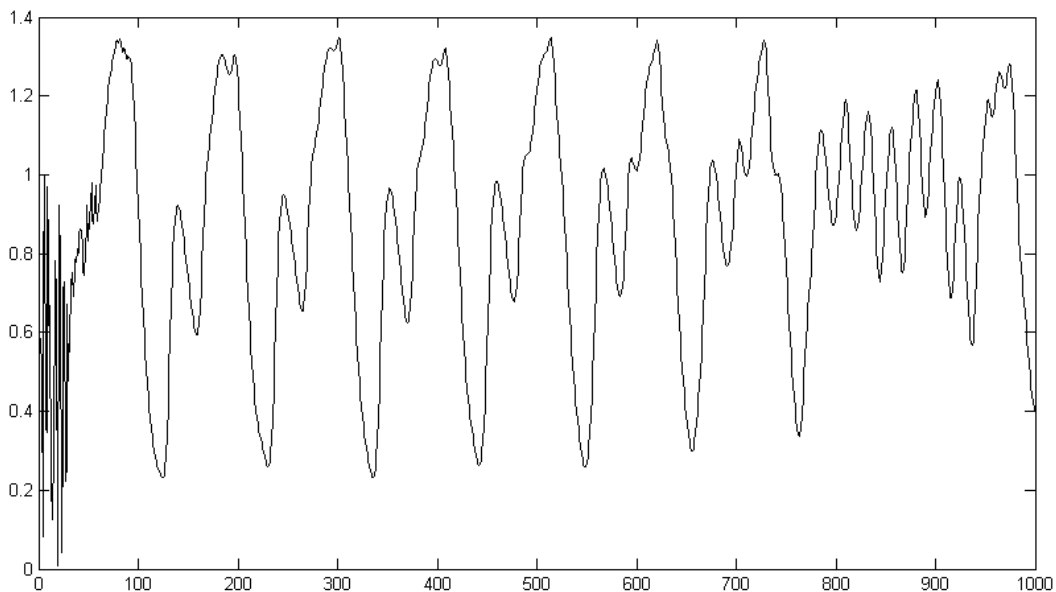


Рисунок 1. График последовательности Маккея-Гласса

В эксперименте использовалась последовательность, состоящая из 5000 элементов такой последовательности. Первые 1000 не рассматривались. 2000 элементов использовались как обучающая выборка, 1000 использовались как тестовая выборка, 1000 элементов использовались как валидационную выборку. Обучающая выборка использовалась для оценки значений (6). По тестовой выборке определялась остановка процесса обучения. Валидационной выборке определялась окончательная ошибка равная (11).

$$error = \sum_t (d(t) - y(t))^2 \quad (11)$$

Количество нейронов в скрытом слое во всех рассматриваемых персептронах было равно 5. В результате были получены следующие данные для ошибки на валидационной выборке, приведенные в таблице 1.

Таблица 1.

Название таблицы

Тип сети	BP7_8	BP7_8_bord	BP9_10	BP9_10_bord
Значение ошибки	1.1023	0.0188	1.1223	0.0592

Наилучшие результаты были получены для алгоритма обучения с использованием (7) и (8) с ограничением коэффициентов связи w_i . При этом результаты для алгоритмов с ограничениями коэффициентов связи w_i лучше на порядки результатов алгоритмов без ограничения.

Из чего можно сделать вывод, что для предложенного персептрона применим алгоритм обучения обратного распространения ошибки со следующими модификациями: градиент оценивается по (7) и (8), на значения коэффициентов связи накладывается ограничение.

Список литературы:

1. Слядников Е.Е. Физическая модель и ассоциативная память информационной биомакромолекулы // Журнал технической физики. — 2007. — Т. 32. — № 8. — С. 52—59.
2. Слядников Е.Е. О взаимосвязи физических и информационных характеристик в окрестности точки сегнетоэлектрического перехода в системе микротрубочки цитоскелета // Журнал технической физики. — 2009. — Т. 79. — № 7. — С. 1—12.
3. Слядников Е.Е. Микроскопическая модель и фазовая диаграмма дипольной системы микротрубочки цитоскелета при конечных температурах // Журнал технической физики. — 2010. — Т. 80. — № 5. — С. 32—39.

4. Слядников Е.Е. Физические основы, модели представления и распознавания образов в микротрубочке цитоскелета нейрона // Журнал технической физики. — 2011. — Т. 81. — № 12. — С. 1—33.
5. Rumelhart, David E.; Hinton, Geoffrey E., Williams, Ronald J. Learning representations by back-propagating errors // Nature. — 1986. — vol. 323. — № 6088. — P. 533—536.
6. Mackey M.C., L. Glass. Oscillation and chaos in physiological control systems // Science. — 1977. — vol. 197. — № 4300. — P. 287—289.