

УДК 519.7

DOI: 10.20310/1810-0198-2016-21-6-2121-2127

ИДЕНТИФИКАЦИЯ ОКРЕСТНОСТНОЙ МОДЕЛИ НЕЙРОННОЙ СЕТИ НА ОСНОВЕ ЖАДНОГО, «ПОЛУЖАДНОГО» АЛГОРИТМОВ И АЛГОРИТМА КАЧМАЖА

© А. М. Шмырин, И. А. Седых, В. В. Семина

Липецкий государственный технический университет
398600, Российская Федерация, г. Липецк, ул. Московская, 30
E-mail: amsh@lipetsk.ru

Рассмотрены жадный, «полужадный» алгоритмы и алгоритм Качмажа для идентификации окрестностной модели нейронной сети и приведены примеры их реализации.

Ключевые слова: окрестностная модель нейронной сети; параметрическая и структурная идентификация; жадный алгоритм; «полужадный» алгоритм; алгоритм Качмажа

1. Введение

В настоящее время теория окрестностных систем является универсальным средством моделирования огромного класса дискретных распределенных систем: стационарных и динамических, четких и нечетких, с линейными и нелинейными связями. Окрестностные модели развивают общие подходы теории систем и теории управления и обобщают такие традиционные дискретные модели, как конечные и клеточные автоматы, сети Петри, разностные уравнения и т. д. [1-2].

Покажем, что нейронную сеть также можно представить в виде окрестностной модели. Рассмотрим несколько алгоритмов идентификации окрестностной модели нейронной сети, в частности алгоритм обратного распространения ошибки, жадный и «полужадный» алгоритмы, алгоритм Качмажа.

2. Окрестностная модель нейронной сети

Нейронную сеть [3-4] можно представить в виде окрестностной модели. Пусть нейронная сеть состоит из m слоев, в каждом j -ом слое ($j = 1, \dots, m$) находится n_j нейронов. Обозначим как $i^{(j)}$ – i -ый нейрон j -го слоя ($i = 1, \dots, n_j$).

Введем следующие определения. Окрестностью нейрона $i^{(j)}$ по входному воздействию $O_v[i^{(j)}]$ назовем группу нейронов предыдущего ($j - 1$)-го слоя, с которым связан данный нейрон, включая сам нейрон. Очевидно, что окрестности по входному воздействию нейронов входного слоя совпадают с самими нейронами, т. е. $O_v[i^{(j)}] = i^{(1)}$, $i = 1, \dots, n_1$.

Окрестности по входному воздействию нейронов входного слоя назовем окрестностями первого уровня. Окрестности нейронов по входному воздействию, находящихся на n -ом слое (скрытом или выходном) нейронной сети, назовем окрестностями n -го уровня. Заметим, что окрестности нейронов могут пересекаться или совпадать (исключая сам нейрон).

Окрестность нейрона по состоянию совпадает с самим нейроном: $O_x[i^{(j)}] = i^{(1)}$, где j ($j = 1, \dots, m$) – номер слоя; i ($i = 1, \dots, n_j$) – номер нейрона в слое.

Рассмотрим представление нейронной сети в виде окрестностной модели:

$$f \left(\sum_{\alpha \in O_v[a^j]} w_v[\alpha^{j-1}, a^j] v[\alpha^{j-1}] \right) = y[a^j] \quad (1)$$

где $v[\alpha^j]$ – входное воздействие в нейрон α j -го слоя ($j = 2, \dots, m$) нейронной сети; $y[a^j] \in R$ – выход в нейроне a j -го слоя; $w_v[\alpha^{j-1}, a^j] \in R$ – весовой коэффициент связи нейронов α $j-1$ -го слоя и a j -го слоя; $f: R \rightarrow R$ – функция активации.

В случае линейной функции активации для нейронной сети получаем симметричную линейную окрестностную систему:

$$\sum_{\alpha \in O_v[a^j]} w_v[\alpha^{j-1}, a^j] v[\alpha^{j-1}] = y[a^j] \quad (2)$$

3. Жадный алгоритм идентификации окрестностной модели нейронной сети

Жадный алгоритм заключается в принятии локально оптимальных решений на каждом этапе, допуская, что конечное решение также окажется оптимальным [5].

Рассмотрим жадный алгоритм структурной и параметрической идентификации окрестностной модели нейронной сети.

1. Вычислить по формуле (1) выходы $y_i^j = y[a^j]$ всех нейронов a_i^j j -го слоя ($i = 1, \dots, n_j, j = 1, \dots, m$).
2. Вычислить квадратичные ошибки окрестностной модели нейронной сети для каждого нейрона a_i^j по формуле: $(y_i^j - d_i^j)^2$, где $i = 1, \dots, n, j = 1, \dots, m$; y_i^j – выходной сигнал нейрона a_i^j , полученный по модели; d_i^j – желаемый выходной сигнал.
3. Упорядочить в порядке возрастания для каждого j -го слоя все разности, полученные в пункте 2.
4. Выбрать в каждом j -ом слое k_j нейронов с минимальными ошибками (k_j задаются). Весовые коэффициенты связей с остальными нейронами считать равными нулю.
5. Настроить неполносвязную сеть методом обратного распространения ошибки.
6. Перейти к пункту 1.

4. «Полужадный» алгоритм идентификации окрестностной модели нейронной сети

Рассмотрим модификацию рассмотренного выше жадного алгоритма структурной и параметрической идентификации окрестностной модели нейронной сети, которую назовем полужадным алгоритмом.

1. Вычислить по формуле (1) выходы $y_i^j = y[a^j]$ всех нейронов a_i^j j -го слоя ($i = 1, \dots, n_j, j = 1, \dots, m$).
2. Вычислить квадратичные ошибки окрестностной модели нейронной сети для каждого нейрона a_i^j по формуле: $(y_i^j - d_i^j)^2$, где $i = 1, \dots, n, j = 1, \dots, m$; y_i^j – выходной сигнал нейрона a_i^j , полученный по модели; d_i^j – желаемый выходной сигнал.
3. Упорядочить в порядке возрастания для каждого j -го слоя все разности, полученные в пункте 2.
4. Выбрать в каждом j -ом слое 2 нейрона с минимальными ошибками. Соответствующие им ошибки обозначить E_1^j и E_2^j соответственно.
5. Добавлять по одному нейрону в каждом слое до тех пор, пока $E_i^j - E_{i-1}^j \rightarrow \min, i = 2, \dots, n_j, j = 1, \dots, m$. Весовые коэффициенты связей с остальными нейронами считать равными нулю.
6. Настроить неполносвязную сеть методом обратного распространения ошибки.
7. Перейти к пункту 1.

5. Примеры идентификации на основе алгоритма обратного распространения ошибки, жадного и «полужадного» алгоритмов

Приведем пример идентификации окрестностной модели нейронной сети с применением алгоритма обратного распространения ошибки, жадного и «полужадного» алгоритмов. Сравним полученные результаты. Рассмотрим многослойную полносвязную нейронную сеть прямого распространения. Пусть исходная нейронная сеть состоит из трех слоев, включая входной и выходной, причем:

- 1) входной слой содержит 10 нейронов;
- 2) скрытый слой – 10 нейронов;
- 3) выходной слой – 1 нейрон.

Рассмотрим обучение окрестностной модели нейронной сети NN на основе данных, равномерно распределенных на интервале $[0,1]$. Будем производить аппроксимацию функции вида:

$$f(v) = v_1^{10} + v_2^9 + v_3^8 + v_4^7 + v_5^6 + v_6^5 + v_7^4 + v_8^3 + v_9^2 + v_{10}, \quad (3)$$

где v_i ($i = 1, \dots, 10$) – входные сигналы нейронов входного слоя.

В ходе алгоритма идентификации окрестностной модели нейронной сети с использованием жадного и «полужадного» алгоритмов получен скрытый слой, состоящий из 6 нейронов. Результаты обучения представлены в табл. 1.

Таблица 1

Ошибка обучения окрестностной модели нейронной сети
на шаге k ($k = 50 \dots 250$ с шагом 50)

Шаг	Алгоритм обратного распространения ошибки	Жадный алгоритм	«Полужадный» алгоритм
50	0.6857	0.1878	0.0705
100	0.63	0.1276	0.016
150	0.0181	0.0167	0.0043
200	0.0071	0.0071	0.001
250	0.0056	0.0266	0.0004

Таким образом, сравнивая результаты применения алгоритма обратного распространения ошибки, жадного и «полужадного» алгоритмов, можно заметить, что:

- 1) «полужадный» жадный алгоритм сходится быстрее, чем остальные;
- 2) при использовании жадного и «полужадного» алгоритмов упрощается структура сети;
- 3) ошибка обучения при применении полужадного алгоритма меньше, чем при использовании алгоритма обратного распространения ошибки и жадного алгоритма.

6. Примеры применения алгоритмов Качмажа и обратного распространения ошибки

Алгоритм, описанный Качмажем в 1937 г., представляет собой классический метод решения системы линейных уравнений, относительно неизвестных значений весовых коэффициентов. Обучение сети описывается итерационной последовательностью по следующему алгоритму:

1. Инициализация весовых коэффициентов w_{ij} ;
2. Предъявление параметров обучения x_{ij} из обучающего множества;
3. Выход нейронной сети $y_i = \sum_{j=1}^n x_{ij}w_{ij}$;
4. Ошибка идентификации системы $\Delta d_i = d_i - y_i$, где d_i – желаемый выход нейрона i ;

5. Вычисление весовых коэффициентов по формуле:

$$w_{ij} = w_{ij} + \frac{\Delta d_i x_{ij}}{\sum_{j=1}^n x_{ij}^2}. \quad (4)$$

Приведем пример параметрической идентификации окрестностной модели нейронной сети с помощью алгоритмов Качмажа и обратного распространения ошибки [7]. Рассмотрим многослойную полносвязную нейронную сеть прямого распространения, состоящую из трех слоев, включая входной и выходной, причем:

- 1) входной слой содержит 3 нейрона;
- 2) скрытый слой – 3 нейрона;
- 3) выходной слой – 1 нейрон.

Рассмотрим обучение окрестностной модели нейронной сети на основе данных, равномерно распределенных на интервале $[0,1]$ и случайно искаженных шумом. Будем производить аппроксимацию линейной функции вида:

$$f(v) = c_1 v_1 + c_2 v_2 + c_3 v_3, \quad (5)$$

где v_i ($i = 1, 2, 3$) – входные сигналы нейронов входного слоя; c_i – заданные константы.

Сравним результаты обучения окрестностной модели нейронной сети, полученные с помощью алгоритма Качмажа и алгоритма обратного распространения ошибки. Результаты для обучения за 150 шагов представлены на рис. 1-2.

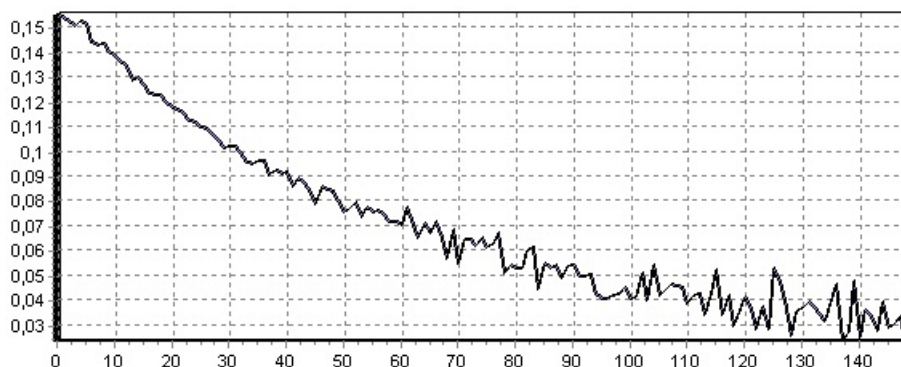


Рис. 1. График ошибки обучения окрестностной модели нейронной сети по алгоритму обратного распространения ошибки в зависимости от шага

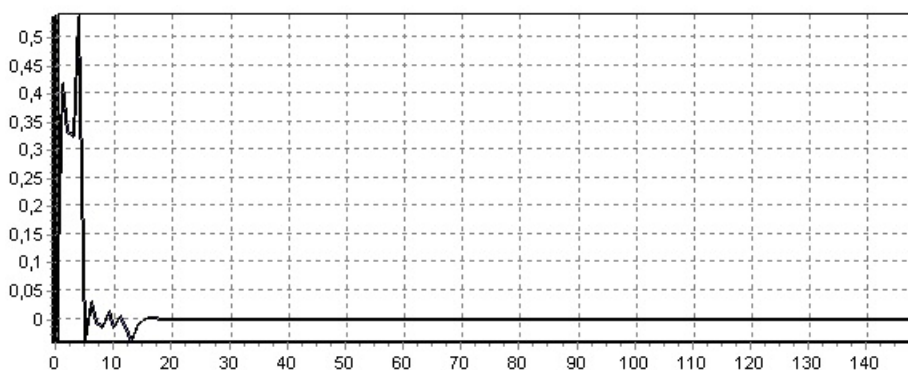


Рис. 2. График ошибки обучения окрестностной модели нейронной сети по алгоритму Качмажа в зависимости от шага

Таким образом, на одинаковых данных ошибка аппроксимации линейной функции с применением алгоритма Качмажа меньше, чем с применением алгоритма обратного распространения. Алгоритм Качмажа с вычислительной точки зрения реализуется значительно проще, чем алгоритм обратного распространения. Кроме того, следует отметить быструю сходимость алгоритма Качмажа.

7. Заключение

В работе показано, что нейронную сеть можно представить в виде окрестностной модели. Рассмотрены жадный, «полужадный» алгоритмы и алгоритм Качмажа для идентификации окрестностной модели нейронной сети.

Приведены примеры идентификации окрестностной модели нейронной сети с применением алгоритма обратного распространения ошибки, жадного и «полужадного» алгоритмов. При сравнении результатов применения алгоритма обратного распространения ошибки, жадного и «полужадного» алгоритмов сделан вывод, что «полужадный» алгоритм сходится быстрее, чем остальные, при этом при использовании жадного и «полужадного» алгоритмов упрощается структура сети. Ошибка обучения при применении полужадного алгоритма меньше, чем при использовании алгоритма обратного распространения ошибки и жадного алгоритма.

Рассмотрен пример параметрической идентификации окрестностной модели нейронной сети с помощью алгоритмов Качмажа и обратного распространения ошибки. Результаты показывают, что при использовании алгоритма Качмажа ошибка аппроксимации меньше по сравнению с алгоритмом обратного распространения. Кроме того, алгоритм Качмажа реализуется проще и сходится быстрее алгоритма обратного распространения ошибки.

СПИСОК ЛИТЕРАТУРЫ

1. Блюмин С.Л., Шмырин А.М., Шмырина О.А. Билинейные окрестностные системы. Липецк; ЛЭГИ, 2006. 131 с.
2. Блюмин С.Л., Шмырин А.М. Окрестностные системы. Липецк; ЛЭГИ, 2005. 132 с.
3. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия – Телеком, 2002. 382 с.
4. Корниенко Н.А. Моделирование неполносвязных нейронных сетей окрестностными системами / Н.А. Корниенко, А.М. Шмырин, И.А. Седых, Т.А. Шмырина // Управление развитием крупномасштабных систем (MLSD-2010): материалы четвертой междунар. конф. Т. 1. М.: ИПУ РАН, 2010. С. 325–327.
5. Шмырин А.М., Седых И.А. Дискретные модели в классе окрестностных систем // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2012. Т. 17. Вып. 3. С. 867–871.
6. Шмырин А.М., Седых И.А. Обобщение дискретных моделей окрестностными системами / А.М. Шмырин, И.А. Седых, Н.А. Корниенко, Т.А. Шмырина // Технические и программные средства систем управления, контроля и измерения (УКИ 10): материалы конф. с международным участием. М.: ИПУ РАН, 2010. С. 207–208.
7. Кормен Т.Х. Алгоритмы: построение и анализ / Томас Х. Кормен, Чарльз И. Лейзерсон, Рональд Л. Ривест, Клиффорд Штайн. М.: «Вильямс», 2006. 1296 с.

БЛАГОДАРНОСТИ: Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 16-07-00854).

Поступила в редакцию 10 октября 2016 г.

Шмырин Анатолий Михайлович, Липецкий государственный технический университет, г. Липецк, Российская Федерация, доктор технических наук, профессор, зав. кафедрой высшей математики, e-mail: amsh@lipetsk.ru

Седых Ирина Александровна, Липецкий государственный технический университет, г. Липецк, Российская Федерация, кандидат физико-математических наук, доцент, доцент кафедры высшей математики, e-mail: sedykh-irina@yandex.ru

Семина Валерия Владимировна, Липецкий государственный технический университет, г. Липецк, Российская Федерация, ассистент кафедры высшей математики, e-mail: pravilnik@mail.ru

UDC 519.854

DOI: 10.20310/1810-0198-2016-21-6-2121-2127

IDENTIFICATION NEIGHBORHOOD NEURAL NETWORK MODEL BASED GREEDY, «SEMIGREEDY» ALGORITHMS AND KACZMARZ ALGORITHMS

© A. M. Shmyrin, I. A. Sedykh, V. V. Semina

Lipetsk State Technical University
30 Moskovskaya St., Lipetsk, Russian Federation, 98600
E-mail: amsh@lipetsk.ru

The article deals with greedy, «semigreedy» Kaczmarz algorithms and an algorithm for the identification of the Neighbor neural network model and examples of their implementation.

Key words: neighborhood neural network model; parametric and structural identification; greedy algorithm; «semigreedy» algorithm; Kaczmarz algorithm

REFERENCES

1. Blyumin S.L., Shmyrin A.M., Shmyrina O.A. Bilinejnye okrestnostnye sistemy. Lipetsk, Legi, 2006. 131 p.
2. Blyumin S.L., Shmyrin A.M. Okrestnostnye sistemy. Lipetsk, Legi, 2005. 132 p.
3. Kruglov V.V. Borisov V.V. Iskusstvennye nejronnye seti: teoriya i praktika. M.: Goryachaya liniya telekom, 2002. 382 p.
4. Kornienko N.A. Modelirovanie nepolnosvyaznyh nejronnyh setej okrestnostnymi sistemami / N.A. Kornienko, A.M. Shmyrin. I.A. Sedyh, T.A. Shmyrina // Upravlenie razvitiem krupnomasshtabnyh sistem MIsd 2010: materialy chetyvortoj mezhdunar konf. 2010. T. 1. M. IPU RAN. P. 325–327.
5. Shmyrin.A.M., Sedyh I.A. Diskretnye modeli v klasse okrestnostnyh sistem // Vestnik Tambovskogo universiteta. Seriya Estestvennye i tekhnicheskie nauki – Tambov University Review. Series: Natural and Technical Sciences. 2012. T. 17. Vyp. 3. P. 867–871.
6. Shmyrin.A.M., Sedyh I.A. Obobshchenie diskretnykh modelej okrestnostnymi sistemami / A.M. Shmyrin. I.A. Sedyh, N.A. Kornienko, T.A. Shmyrina // Tekhnicheskie i programmnye sredstva sistem upravleniya kontrolya i izmereniya (UKI 10): materialy conf. s mezhdunarodnym uchastiem. M. IPU RAN, 2010. P. 207–208.
7. Kormen T. H. Algoritm:y postroenie i analiz / Tomas H. Kormen, Charlz I. Lejzerson, Ronald L. Rivest, Klifford Shtajn. M: Vilyams, 2006. 1296 p.

ACKNOWLEDGEMENTS: The work is partially supported by the Russian Fund for Basic Research (project № 16-07-00854).

Received 10 October 2016

Shmyrin Anatoliy Mihaylovich, Lipetsk State Technical University, Lipetsk, the Russian Federation, Doctor of Techniques, Professor, the Head of the High Mathematics Department, e-mail: amsh@lipetsk.ru

Sedich Irina Aleksandrovna, Lipetsk State Technical University, Lipetsk, the Russian Federation, Candidate of Physics and Mathematics, Associated Professor of the High Mathematics Department, e-mail: sedykh-irina@yandex.ru

Semina Valeriya Vladimirovna, Lipetsk State Technical University, Lipetsk, the Russian Federation, Lecturer of the High Mathematics Department, e-mail: pravilnik@mail.ru

Информация для цитирования:

Шмырин А.М., Седых И.А., Семина В.В. Идентификация окрестностной модели нейронной сети на основе жадного, «полужадного» алгоритмов и алгоритма Качмажа // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2016. Т. 21. Вып. 6. С. 2121-2127. DOI: 10.20310/1810-0198-2016-21-6-2121-2127

Shmyrin A.M., Sedykh I.A., Semina V.V. Identifikatsiya okrestnostnoj modeli nejronnoj seti na osnove zhadnogo, «poluzhadnogo» algoritmov i algoritma Kachmazha [Identification neighborhood neural network model based greedy, «semigreedy» algorithms and Kaczmarz algorithms]. *Vestnik Tambovskogo universiteta. Seriya Estestvennye i tekhnicheskie nauki – Tambov University Review. Series: Natural and Technical Sciences*, 2016, vol. 21, no. 6, pp. 2121-2127. DOI: 10.20310/1810-0198-2016-21-6-2121-2127 (In Russian)